DOCUMENT RESUME

ED 320 857 SP 032 101

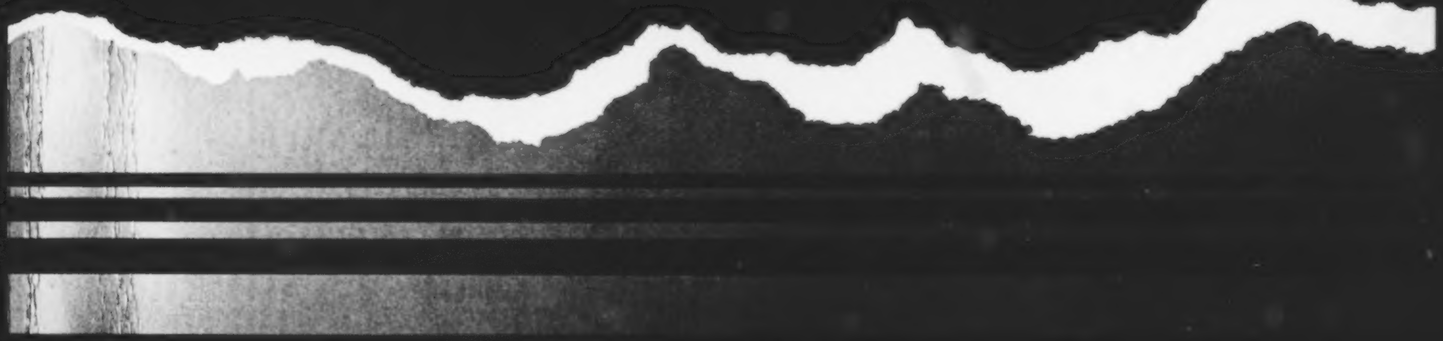
TITLE	Physical Fitness Promotion. Program Evaluation Handbook.
INSTITUTION	IOX Assessment Associates, Culver City, CA.
SPONS AGENCY	Centers for Disease Control (DHHS/PHS), Atlanta, GA.; Public Health Service (DHHS), Rockville, MD. Office of Disease Prevention and Health Promotion.
PUB DATE	88
NOTE	213p.; For the other titles in this series, see SP 032 099, SP 032 102-105, and SP 032 107.
AVAILABLE FROM	IOX Assessment Associates, 5420 McConnell Ave., Los Angeles, CA 90066 ($17.95).
PUB TYPE	Guides - Non-Classroom Use (055)
EDRS PRICE	MF01/PC09 Plus Postage.
DESCRIPTORS	*Evaluation Criteria; Evaluation Methods; Exercise; Health Education; *Physical Fitness; *Program Evaluation; *Psychometrics

ABSTRACT

Intended as a resource for individuals wishing to evaluate physical fitness programs, this handbook, one of a series of seven, provides a collection of measuring devices that can improve the quality of such evaluations. Chapter 1 introduces the handbook's contents and outlines evaluation related issues specific to physical fitness programs. Chapter 2 introduces the key operations involved in program evaluation, emphasizing the role of assessment instruments in the gathering of information needed for defensible evaluations. Chapter 3 treats physical fitness measures and provides checklists and facts on regular exercise. The fourth chapter, on physiological measures, contains the measuring tools designed to be used in the evaluation and design of physical fitness programs. These measures deal with behavior, knowledge, skills, and affective outcomes. Each measure is introduced by a brief description of the purpose of the assessment instrument, as well as procedures for administering, scoring, and analyzing the resulting data. Chapter 5 describes how technical appraisals of the handbook's measures can be carried out. Appendixes contain amplified content descriptions for updating the various measures, an explanation of informed consent procedures, and an annotated bibliography. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED320857



SP032 161

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

IOX Assessment
Associates

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

PROGRAM EVALUATION HANDBOOK

PHYSICAL FITNESS PROMOTION

Prepared for

The Center for Health Promotion and Education
United States Centers for Disease Control

The Office of Disease Prevention and Health Promotion
Office of the Assistant Secretary for Health
United States Department of Health and Human Services

by

IOX Assessment Associates
P.O. Box 24095
Los Angeles, CA 90024-0095

1988

TABLE OF CONTENTS

Preface vi

1 **A Resource for the Evaluation of Physical Fitness Promotion Programs** 1

An Evidence-Oriented Era 3
Measurement and Program Design 3
What the Handbook Contains 4
How to Use the Handbook 5
Technical Quality of the Handbook's Measures 7
Physical Fitness Promotion 7
Physical Fitness Terminology 8
References 9

2 **Essentials of Program Evaluation for Health Educators** 11

Focusing the Evaluation 13
Rights of Participants .16
Selecting Appropriate Measures 17
When to Administer Measures 19
Data-Gathering Design Options 21
Sampling Considerations for Data Collection 28
Data Analysis 30
Reporting Results 30
Reprise 31

3 **Physical Fitness Promotion Measures** 33

Overview of Measures .35
Weekly Activities Index 37
Injury Checklist 43
Facts About Exercise .46
Exercise Facts 52
Planning a Safe Exercise Program 56
Planning to Exercise Safely 62
Selecting an Exercise Program 66

Preventing and Caring for Injuries 80
Exercising Safely 90
Ideas About Decisions 96
Effects of Exercise 99
Beliefs About Exercise 101
Exercising Regularly 103
Intention to Exercise 106
Attitude Toward Work 109

4 **Physiological Measures 111**

Considerations for Physiological Testing 113
Overview of Physiological Measures 115
Physical Activity Readiness Questionnaire (PAR-Q) 117
Bruce Treadmill Test 119
Bicycle Ergometer 123
3-Minute Step Test 133
Cooper's 1.5-Mile or 12-Minute Run/Walk 137
Distance Run 140
Jackson and Pollock's Skinfold Measures 143
Sum of Skinfold Fat 151
Waist Circumference 156
Buttocks (Hip) Circumference 158
Push-ups (Canadian Fitness Test) 160
Fitnessgram® Flexed-arm Hang 163
Sit-ups 166
Modified Sit-ups 169
Sit and Reach Test 172
AAHPERD Sit and Reach Test 176
References 180

5 **Locally Conducted Psychometric Studies 181**

Determining the Technical Quality of Measuring Devices 183
Categories of Validity Evidence 183
Types of Reliability 187
Groups and Individuals 188

Appendices **191**

A. Amplified Content Descriptors 193
  Planning to Exercise Safely/Planning a Safe Exercise Program 193
  Facts About Exercise/Exercise Facts 197
  Physical Fitness Promotion Content Bibliography 201

B. Informed Consent Procedures 202

C. Annotated Evaluation Bibliography 203

Preface

In recent years, health educators have increasingly recognized that systematic evaluation can help them appraise and improve their programs. For this potential to be realized, however, effective mechanisms for gathering relevant data are required. In the past, critical information about a program's effects was not collected in some instances because suitable measures for gauging those effects were lacking. The purpose of this handbook is to rectify, at least in part, this deficiency in the evaluation of health education programs dealing with physical fitness promotion.

This book is one of seven health education evaluation handbooks resulting from a project jointly initiated in 1980 by the United States Centers for Disease Control (CDC) and the Office of Disease Prevention and Health Promotion (ODPHP) of the Office of the Assistant Secretary for Health. The handbook is not intended to be prescriptive or all-inclusive. Those who evaluate physical fitness promotion programs should regard the handbook as only a resource, that is, a collection of assessment tools that may be of use in program evaluation. The extent to which the handbook will actually be useful depends chiefly on the extent to which it contains assessment tools that correspond to the evaluation needs of a particular physical fitness promotion program.

Handbook Development

This handbook has been created by IOX Assessment Associates (IOX), selected competitively on the basis of responses to a governmentally issued request for proposals. IOX was to collect and develop program evaluation measures for critical behavior, knowledge, skill, and affective outcomes in the area of physical fitness. Three panels of experts played prominent roles in the creation of this handbook. A **Handbook-Development Panel**, consisting of six experts familiar with physical fitness promotion programs or their evaluation, guided the initial development of the handbook. The Handbook-Development Panel identified important outcomes for physical fitness promotion programs. IOX staff, drawing on the advice of panelists, then developed assessment instruments to assess panel-identified program outcomes. The names and affiliations of the Physical Fitness Promotion Handbook-Development Panelists are provided on the following page.

Handbook-Development Panel

Dr. Sharon Dorfman
Division of Health Education
Johns Hopkins University
Baltimore, Maryland

Dr. William L. Haskell
School of Medicine
Stanford University
Stanford, California

Dr. Steven Havas
Bureau of Health Promotion
and Disease Prevention
Connecticut Department of Health
Hartford, Connecticut

Dr. Ash Hayes
The President's Council on
Physical Fitness and Sports
Washington, D.C.

Dr. Wells Hively
Central Midwestern Regional
Education Laboratory
St. Louis, Missouri

Dr. Ralph Paffenbarger
School of Medicine
Stanford University
Stanford, California

The Handbook-Development Panel met at the beginning of the project in order to isolate the chief outcomes that physical fitness programs could reasonably be expected to promote. Preliminary statements reflecting these outcomes were identified by the panelists. These preliminary outcome statements were refined by IOX staff and mailed to the panelists and other interested specialists, all of whom rated the importance of each statement. The list of high-priority outcomes that resulted was used to guide the selection and development of the original handbook's measures.

All newly developed measures were mailed to the panelists for review. In addition, all of these measures were tried out with small groups of respondents. The measures were revised based on the informal tryouts and the panelists' review comments. All of the new measures were also reviewed by IOX staff in an effort to eliminate any potential ethnic, gender, religious, or socioeconomic bias.

A completed version of the physical fitness promotion handbook was delivered to the government in 1983. Several thousand copies of the handbook were released by CDC and ODPHP to health educators throughout the nation.

Handbook Revision

Subsequent to the initial distribution of the handbook, CDC issued, in concert with ODPHP, a second request for proposals which led to the comprehensive revision of the existing physical fitness promotion handbook. To guide the review and revision of the physical fitness handbook, a **Handbook-Revision Panel** was constituted. Members of the panel were selected because of their dual expertise in (a) the field of physical fitness and (b) measurement of the outcomes sought by physical fitness promotion programs. Members of the Handbook-Revision Panel and their affiliations are listed on the following page.

Handbook-Revision Panel

Dr. Steven Blair
Institute for Aerobic Research
Dallas, Texas

Dr. Ronald La Porte
University of Pittsburgh
Pittsburgh, Pennsylvania

Dr. William Zuti and Associates
Chicago, Illinois

Dr. D.W. Edington
University of Michigan
Ann Arbor, Michigan

Dr. Jonathan E. Fielding
U.S. Corporate Health Management
and University of California
Los Angeles, California

Dr. Glen G. Gilbert
Office of Disease Prevention
and Health Promotion
Washington, D.C.

Dr. William Haskell
Stanford University
Stanford, California

Dr. Michael Pollock
Mount Sinai Medical Center
Milwaukee, Illinois

Dr. Peter Cortese
California State University
Long Beach, California

Dr. Donald Iverson
University of Colorado
Denver, Colorado

Dr. Kenneth Powell
Centers for Disease Control
Atlanta, Georgia

The Handbook-Revision Panel met on two occasions. In these meetings, panelists reviewed the contents of the initial version of the physical fitness handbook, particularly its measures, and suggested deletions, modifications, or additions. Panelists also provided guidance regarding ways of making the handbook more usable to practitioners. During both of these meetings, the panelists were attentive to the accuracy of the handbook's contents. Considerable content, in the measures as well as the introductory materials, was revised or deleted on the basis of panelists' suggestions.

Overall Guidance

A third panel, the **Project Advisory Panel**, provided overall guidance to IOX staff during the final three years of the project. These individuals offered technical counsel and strategic advice during the revision of all handbooks. Members and affiliations of the Project Advisory Panel are listed on the following page.

Project Advisory Panel

Dr. Peter A. Cortese
California State University
Long Beach, California

Dr. William L. Haskell
Stanford University
Stanford, California

Dr. Lawrence W. Green
Henry J. Kaiser Family Foundation
Menlo Park, California

Dr. Jonathan E. Fielding
U.S. Corporate Health Management
and University of California
Los Angeles, California

Acknowledgments

The project that led to the creation of this handbook was funded by the CDC and ODPHP. Dr. Walter J. Gunn of CDC conceptualized the project and supplied technical guidance throughout its first phase. During this time, Dr. Diane Orenstein of CDC as well as Dr. Donald Iverson and Dr. Patricia Mullen, both of ODPHP, served as project officers.

During the project's second phase, that is, the revision of the handbook, Dr. Orenstein of CDC continued to serve as project officer along with Dr. Joel Kavet, Dr. Gregory Christenson and Mr. James Harrell of ODPHP.

As the handbook progressed, numerous health educators throughout the nation offered their insights regarding the handbook's contents. Without their expert assistance, development of this volume would have been impossible.

IOX Assessment Associates
July, 1988

CHAPTER ONE

A Resource for the Evaluation of Physical Fitness Promotion Programs

A Resource for the Evaluation of Physical Fitness Promotion Programs

This handbook is intended to help those individuals who wish to evaluate health education programs dealing with physical fitness. More specifically, the handbook provides a series of measuring devices that, if selected and used judiciously, can improve the quality of such evaluations. As a consequence, not only will the technical quality of the program evaluation be improved, but any program-related decisions based on the evaluation's results are apt to be more defensible.

An Evidence-Oriented Era

In recent years, educators have experienced substantially increased pressures to produce evidence that their programs are functioning effectively. In contrast to an earlier era when it was widely thought that most educational programs were worth the money they cost, today's educators find that they are constantly called on to justify the effectiveness of their programs.

The kinds of evidence that health educators have been required to assemble regarding program effectiveness have, almost without exception, involved the use of various kinds of assessment instruments. Consonant with that requirement, this handbook contains numerous tests and inventories designed to secure the evidence needed to judge the effectiveness of physical fitness promotion programs. The handbook's measuring instruments were created specifically to assess important goals of the most common types of physical fitness promotion programs offered for adults (in industrial or clinical settings) and for children (in school-related programs).

The handbook, accordingly, makes available to those who operate physical fitness promotion programs the assessment tools by which the effectiveness of such programs can be determined. The evidence of program effectiveness currently being demanded of physical fitness promotion personnel can, therefore, be provided by appropriate use of the handbook's assessment instruments. Moreover, as will be indicated shortly, appropriate use of the handbook's numerous assessment devices can substantially improve the *design* of physical fitness promotion programs.

Measurement and Program Design

Historically, assessment devices have been thought of as instruments to be used *after* a program was concluded. Teachers, for example, have traditionally administered tests *after* instruction was over in order to grade students. However, even though assessment instruments have often been post-instruction creations of instructors, such instruments can make important – often overlooked – contributions to the original design of an instructional program. Properly developed assessment tools, in fact, can contribute to program design in two significant ways.

First, because assessment instruments are typically intended to measure outcomes of interest, such assessment instruments provide program personnel with a range of potential outcomes. An increased range of possible program outcomes generally leads to the *selection of more defensible outcomes* for health education programs. To illustrate, there may be an assessment instrument dealing with an attitudinal dimension that, were it not for the measuring instrument's availability, might have been overlooked by the program staff. Stimulated by the assessment tool's availability, however, the program staff can add the attitudinal dimension to the program's targeted outcomes.

A second program-design dividend of properly constructed assessment tools is that they *clarify intended program outcomes* and, thereby, make possible the provision of more on-target program activities than would have been the case had such clarification not been present. To illustrate, suppose that program personnel intend to feature in their evaluation an assessment device focused on the knowledge of the effects of exercise. By becoming familiar with the composition of that assessment tool, the program staff can be sure to incorporate critical facts about those effects in their instructional program. Provision of appropriate instructional practice for participants need not reflect "teaching to the test" in the negative sense that instructors coach students for specific test items. Instead, providing relevant knowledge so that program participants attain the program's intended outcomes constitutes an efficient and effective, research-supported form of instruction.

To review, then, the measuring instruments provided in this handbook are intended to assist those who design and those who evaluate physical fitness promotion programs. With respect to program evaluation, the measures will yield evidence by which to improve programs as well as determine program effectiveness. With respect to program design, the measures provide a menu of potential program options and, once having been selected, enhanced clarity regarding the nature of the outcome(s) sought.

What the Handbook Contains

There are several key ingredients in this handbook. It should, therefore, prove helpful to readers if the handbook's major sections are presented. Briefly, then, here is a description of the handbook's major components:

Introductory information. In Chapter One, an introduction to the handbook is provided. Because the handbook is intended to be used with physical fitness promotion programs, the chapter concludes with a brief discussion of evaluation-related issues specific to health education programs dealing with physical fitness promotion.

Program evaluation essentials. Although a number of people who use this handbook will already be familiar with the nature of program evaluation, many handbook users will not be well versed in the conduct of program evaluations. Accordingly, in Chapter Two, an introduction is provided to the key operations involved in program evaluation. Although space limitations preclude a detailed exposition of all aspects of program evaluation, emphasis is given to the role that assessment instruments play in the gathering of information needed for defensible evaluations.

Assessment instruments. Chapter Three contains one of the handbook's most important components, namely, the measuring tools designed to be used in the evaluation and design of physical fitness promotion programs. These measures deal with behavior, knowledge, skill, and affective outcomes. *Behavior* measures focus on actual behaviors of program participants. *Knowledge* measures are concerned with participant mastery of a defined set of information. *Skill* measures deal with cognitive, that is, intellectual, competencies to be mastered by program participants. Finally, *affective* measures assess participants' attitudes and values.

In addition to the newly developed measures in Chapter Three, a selection of extant fitness-testing measures are provided in Chapter Four. These physiological measures assess cardiorespiratory function, body composition, muscular strength and endurance, and lower trunk flexibility.

Each measure is introduced by a brief description of the purpose of the assessment instrument, as well as procedures for administering, scoring, and analyzing the resulting data. All measures have been provided on detachable pages. At the beginning of both Chapters Three and Four, an overview of the chapter's measures is provided to facilitate the selection of measures.

Local measure appraisal. Although the measures contained in Chapter Three have been created with considerable care and were pilot tested in small-scale tryouts, the measures have not yet been subjected to a formal empirical appraisal of their technical adequacy. Thus, in Chapter Five, a description is provided of how such technical appraisals of the handbook's measures can be carried out.

Annotated bibliography. Because evaluators and designers of physical fitness promotion programs may wish to consult additional sources regarding program design and program evaluation, an annotated bibliography is provided in Appendix C to facilitate the handbook user's selection of such materials.

Amplified content descriptors. The information eligible for inclusion in the knowledge measures is provided in Appendix A as amplified content descriptors. Additional content that can be used for the generation of new items is also presented. However, these descriptors are not exhaustive accounts of physical fitness promotion content.

How to Use the Handbook

The particular ways in which the handbook is used will vary from setting to setting and from user to user. For instance, if a handbook user is relatively unfamiliar with the core notions in program evaluation, then a thorough reading of Chapter Two's treatment of program evaluation essentials is warranted. In addition, further reading based on the evaluation-related references included in the annotated bibliography would also seem useful.

For handbook users more familiar with program evaluation, primary attention will probably be focused on the measures in Chapters Three and Four. Although use of the measures will vary from situation to situation, a common four-step usage pattern is depicted in Figure 1.1.

Note that in Step 1, the measures are used to represent a range of potential program objectives. Clearly, an expanded range of options can lead to more appropriate decisions regarding what program objectives to pursue. In Step 2, after the measures for possible program evaluation have been reviewed, one or more measures are selected for use in the evaluation of the program. In Step 3, after the program evaluation measures have been

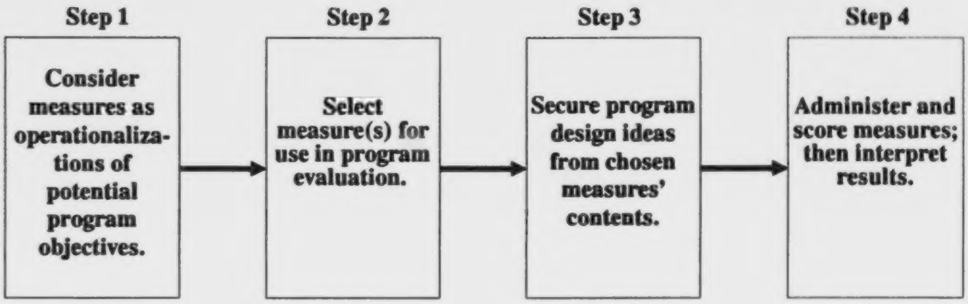

Step 1
Consider measures as operationalizations of potential program objectives.
Step 2
Select measure(s) for use in program evaluation.
Step 3
Secure program design ideas from chosen measures' contents.
Step 4
Administer and score measures; then interpret results.

Figure 1.1: A four-step usage pattern of the handbook's measures

selected, the program staff studies the measures intensively to discern if there are program design implications to be drawn from the measures. In Step 4, the measures are administered using one of the evaluative data-gathering designs described in Chapter Two and scored according to the scoring directions in Chapters Three and Four. Finally, interpretations of the results are made.

It is important to remember that the handbook's measures are to be used for program evaluation, not individual decision making. Thus, if one of the handbook's affective measures was used on a pretest-posttest basis, it is the *aggregation* of scores on the measure that provides us with an indication of the program's effectiveness. The measures were not designed to yield an accurate indication of an *individual* participant's status. Thus, it would be inappropriate to attempt to determine an individual participant's attitudes on the basis of the handbook's measures. The measures are relatively brief instruments designed to be administered without great intrusiveness. When the measures' scores are viewed in the aggregate, the measures can provide data of relevance to program evaluators. The data, however, should *not* be used for determining the status of individuals.

Another point related to use of the handbook's measures concerns the potential *reactivity* of certain measures, that is, the likelihood that if the measure is used *prior* to the program, the experience of completing a measure may cause participants to react differently to the program than had the measure not been administered. Reactivity is more frequently associated with affective measures rather than cognitive measures. Thus, handbook users will need to be alert to the possibility that a given measure, if administered prior to the program, will unduly sensitize participants to an aspect of the program.

To avoid such reactive effects, program personnel may need to divide participants into two subgroups so that only a portion of the participants receive any given potentially reactive measure. Such subgroups would not be given the same reactive measure both before and after the program. Rather, participants should be administered only post-program measures that they had *not* been given prior to the program. Indeed, two potentially reactive measures may be administered simultaneously under the conditions represented in Figure 1.2, where it can be seen that the pre-program performance of certain participants (one-half, for example) serves as a comparison for the post-program performance of other participants. Although a variety of data-gathering designs will be

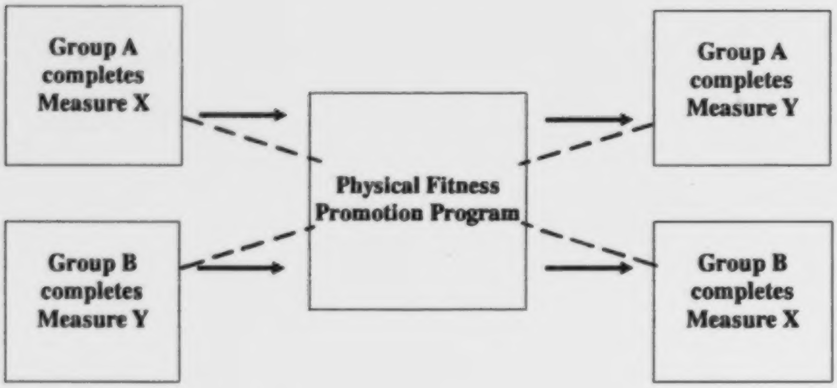

Group A completes Measure X
Group B completes Measure Y
Physical Fitness Promotion Program
Group A completes Measure Y
Group B completes Measure X

Figure 1.2: Using the handbook's measures to avoid reactive effects (Appropriate Comparisons = – – –)

described in Chapter Two, the evaluator should employ care in using the handbook's measures so that they permit reasonable inferences regarding program effectiveness. Potential reactivity of measures should be examined when considering such designs.

Technical Quality of the Handbook's Measures

The measuring instruments to be found in Chapter Three were carefully constructed by an experienced test-development agency according to the guidance of prominent experts in the field of physical fitness education. All of Chapter Three's assessment devices were subjected to small-scale tryouts, revised on the basis of those tryouts, and reviewed by physical fitness promotion specialists.

At the outset of this handbook development project, it had been anticipated that all of the handbook's measuring instruments would be subjected to large-scale field tests so that substantial empirical evidence regarding the technical quality of the measures could be made available to handbook users. Unfortunately, that phase of the project could not be completed.

Thus, handbook users should be cautioned that, although the handbook's measures were developed with great care, there is currently no evidence available by which to ascertain the technical quality of the measures. Therefore, handbook users must exercise caution in the use of Chapter Three's assessment instruments. In Chapter Five, as indicated earlier, a description is presented of the ways in which users of the handbook's measures, if they wish to do so, can carry out local studies regarding the technical quality of the measures that they find most suitable for their use.

Physical Fitness Promotion

The current philosophy of physical fitness promotion has shifted from a motor fitness and athletic performance orientation toward a disease prevention and health promotion perspective. Under the athletic performance philosophy, a "fit" individual is one who possesses a wide variety of athletic abilities including speed, agility, strength, and endurance.

On the other hand, under the disease prevention/health promotion philosophy, health-related fitness is defined as the ability to perform physical activities with vigor and without excessive fatigue (Pate, 1983).

Both the athletic performance and health-related fitness philosophies include the fitness components of cardiorespiratory endurance, muscle strength and endurance, body composition, and flexibility. The athletic performance philosophy also identifies agility, power, speed, and balance as essential elements of fitness. However, only the fitness components of cardiorespiratory endurance, body composition, and neuromuscular function are currently considered important under the disease prevention and health promotion philosophy (AAHPERD, 1980; Pate, 1983).

This handbook includes two separate chapters of measures for the evaluation of fitness promotion programs. The measures in Chapter Three are newly developed test instruments that assess participants' behaviors, knowledge, attitudes, and skills. Consonant with the disease prevention and health promotion philosophy, the physiological measures provided in Chapter Four assess health-related fitness rather than athletic fitness. The measures are grouped into the four categories: cardiorespiratory function, body composition, muscular strength and endurance, and flexibility.

Physical Fitness Terminology

Before proceeding to evaluation considerations, there are several fitness-related terms, which are employed throughout the handbook, that warrant some elaboration.

Frequency refers to the number of times exercise is performed on a regular basis. Exercise should be performed from three to five days per week in order to show cardiorespiratory improvement. Starter programs often restrict activity to 5 - 15 minutes at a time but require participants to engage in an activity two to three times throughout a day (AAHPERD, 1980; Pollock, 1984; Golding, 1984).

Intensity refers to the degree of energy (total kilocaloric expenditure) used during an exercise activity. In order to improve and maintain cardiorespiratory fitness, intensity levels should be above the minimal threshold level, which is 60 percent of a person's maximum heart rate. However, people with low initial levels of fitness (e.g., cardiac patients or individuals who are overweight) can show improvement with programs of less than 60 percent intensity (American Heart Association, 1979; Pollock, 1984; Golding, 1984).

Duration refers to the length of time one is engaged in physical activity. Duration may vary from activity to activity and from day to day. Duration is used in conjunction with intensity to determine the total amount of kilocalories used during an exercise period. The recommended duration depends, in part, on the intensity of the activity performed.

MET refers to the metabolic rate at rest while sitting quietly. METS are commonly used as an alternative method of measuring the intensity level of exercise. One MET is equal to the energy expended or oxygen cost at rest; an activity equivalent to three METS, therefore, requires three times the amount of energy expended at rest. Activities (including household and occupational activities) are divided into four intensity levels (light, moderate, hard, or very hard) according to the amount of energy required. Each activity level has a corresponding MET value. For example, "very hard" activities, such as cross country skiing, are assigned a MET value of 10. The MET is useful because it accounts for differences in body weight without extra calculations. According to the American Heart Association (1979), 110 - 240 minutes of activity (equivalent to 5 METS or higher) a week is needed to maintain a low to moderate risk of developing coronary heart disease (Pollock, 1984; Golding, 1984).

Information on how to use this handbook in evaluating the effectiveness of fitness promotion programs is provided in the next chapter.

References

American Alliance for Health, Physical Education, Recreation, and Dance. (1980). *Health related physical fitness test manual.* Reston, VA: Author.

American Heart Association. (1972). *Exercise testing and training of apparently healthy individuals: A handbook for physicians.* Dallas, TX: Author.

Golding, L.A., Myers, C.R., & Sinning, W.E. (Eds.). (1982). *The Y's way to physical fitness.* Chicago, IL: YMCA of the USA.

Pate, R.R. (1983). A new definition of fitness. *The Physician and Sportsmedicine, 11*(4), 77-82.

Pollock, M.L., Wilmore, J.H., & Fox, S.M. (1984). *Exercise in health and disease: Evaluation and prescription for prevention and rehabilitation.* Philadelphia, PA: W.B. Saunders.

CHAPTER TWO

Essentials of Program Evaluation for Health Educators

Essentials of Program Evaluation for Health Educators

Education programs are intended to help people. Public school programs, for example, are intended to help youngsters acquire the skills and knowledge that they will need as adults. Similarly, health education programs are intended to promote participants' adoption of beneficial health-related behaviors. Yet, even though an education program might have been well intentioned, how do we know that the goals of the program were realized? Moreover, if a program is not meeting its goals, how can the program be made more effective?

Such questions constitute the core of program evaluation. In essence, evaluators want to discover whether a program has worked effectively and, if not, how it can be made more effective. When evaluation is used to improve programs, it can make a significant contribution to the well-being of program participants and, potentially, to the community at large.

In this chapter, the nature of program evaluation will be considered as it relates to health education programs. The following topics will be discussed:

- Focusing the Evaluation
- Rights of Participants
- Selecting Appropriate Measures
- When to Administer Measures
- Data-Gathering Design Options
- Sampling Considerations for Data Collection
- Data Analysis
- Reporting Results

The purpose of this chapter is not to promote a particular evaluation model for health education programs. Rather, the chapter deals with considerations central to any evaluation effort. It is hoped that evaluators* of physical fitness promotion programs will be able to apply the chapter's contents to their endeavors.

Focusing the Evaluation

The results of a program evaluation can be used to improve decisions about programs. Anyone setting out to evaluate a health education program, therefore, should focus the evaluation on the *decisions* that are likely to be made about the program, either while the program is being implemented or when it is concluded. In other words, if evaluators know what decisions are apt to be faced by those who will use the evaluation's results, then information bearing on those decisions should, if possible, be collected during the evaluation. To determine what these decisions are, an evaluator needs to have a clear understanding of the purpose of the program, the specifics of the program, and the individuals or groups who may use the evaluation's results. Focusing the evaluation involves considerations such as (a) the nature and role in the evaluation of program objectives, (b) the summative and formative functions of evaluation, (c) the cost of the program, (d) the extent to which observed changes in participants will also be attributed to the program, and (e) the extent to which program effects will be generalizable to other situations. Each of these considerations is discussed below.

---

* Sometimes a program evaluation will be conducted by an individual not affiliated with the program itself—an individual formally designated as a program evaluator. More frequently, however, an evaluation will be carried out by the personnel who are actually operating the program. Whenever the term "evaluator" is used in this handbook, it will refer both to the evaluator-specialist and to the program staff member serving as evaluator.

Objectives and evaluation. Health education programs are designed to bring about worthwhile effects. Most health education programs, therefore, are organized around some form of program objectives that focus on such intended effects. In general, the more clearly these objectives are stated, the more useful they will be in carrying out an evaluation.

One way of conducting an evaluation is to determine the extent to which a program's objectives have been achieved. Program designers too frequently describe their objectives in such ambiguous, general ways, however, that it is impossible to tell whether such loosely defined objectives have been attained. It is for this reason that it can be beneficial for evaluators to work with program personnel, prior to program implementation, to create program objectives that clearly describe desired post-program participant behaviors.

Another potential pitfall when creating program objectives is the tendency to delineate a set of hyper-detailed objectives. Specificity does not automatically yield utility. Instead, decision makers can become overwhelmed by long lists of low-level, albeit behaviorally stated, objectives. For example, a program objective that participants be able to identify the proper care for a blister is going to lead down a path toward numerous small-scope objectives. Recent thinking regarding instructional objectives suggests that program objectives, *while still measurable,* should focus on larger, more significant types of participant post-program behaviors. A more significant fitness-related objective, for example, might be that participants be able to identify proper care for common exercise-related injuries. Today's health education programs, rather than being organized around 30 minuscule (and, therefore, potentially trivial) objectives, might better be focused on a half-dozen more general, but still measurable, program objectives.

Most evaluators agree, however, that there is substantially more to program evaluation than merely determining whether a program's objectives have been achieved. For example, there may be effects of the program that were not anticipated in the program's stated objectives. Evaluators need to be attentive not only to the effects of a program that were anticipated, but also to any unforeseen program effects.

Summative and formative functions. Summative evaluation addresses the question of whether a program, in its complete and final form, is effective. The decisions associated with the summative evaluation are essentially go/no-go decisions, such as whether to continue a health education program or, perhaps, whether to disseminate the program more widely.

Formative evaluation addresses questions associated with improving a program that is "under development," that is, still modifiable. The decisions associated with formative evaluation focus on ways to improve particular parts of the program. Formative evaluation is an ongoing endeavor conducted as the program is designed, installed, and maintained. Whereas summative evaluation's mission is to provide a final judgment about a program's overall merit, formative evaluation's mission is to bolster a program's quality on a continuing basis. The effective formative evaluator functions less as an external judge and more as a collaborating member of the program team. The formative evaluator's task is to monitor the program so that it can be improved.

Almost all programs are, at least to some degree, modifiable. Hence, only in rare cases do evaluators appraise a health education program in its complete and final form. One such instance might involve a materials-based physical fitness program. For example, if the program were found to be effective via a summative evaluation, a commercial publisher would distribute the program's materials nationally. In most cases, however, health education programs can be modified and improved. Thus, a formative, improvement-oriented evaluation can be carried out for most health education programs.

Cost-analysis considerations. Program evaluators are often so concerned about detecting the effects of programs that they fail to consider the *costs* of those effects. Yet decision makers need information regarding not only the effects of a program, but also the resources required to achieve those results. For this reason, program evaluators should carefully isolate and communicate the relative costs of programs. For example, information should be collected that can show how much Program A costs to produce a given result compared to the cost of Program B to produce a comparable result. Judgments about a program's impact without considerations regarding its costs are potentially superficial. In recent years, there has been much attention to cost-analysis strategies. Although consideration of those procedures is beyond the scope of this handbook, serious evaluators of health education programs would do well to delve more deeply into cost-analysis procedures.*

Attributing observed changes to the program. Characteristically, an evaluation seeks to determine whether individuals have changed as a result of their participation in a program. The key issue is whether pre-program to post-program changes in the status of participants are attributable to the program itself or to other extraneous factors. Examples of extraneous factors are participants' maturation, their familiarity with the measures used in the evaluation, or their reactions to non-program events such as a health-related, mass media campaign. This issue revolves around the evaluator's ability to properly infer that the program itself caused any observed changes in participants. Technically, the degree to which evaluators can validly infer that a program caused a set of observed changes is referred to as the *internal validity* of the evaluation study. Ideally, an evaluation's data-gathering design should help to rule out explanations other than the program itself for observed changes. (Data-gathering design options are discussed later in this chapter.) If evaluators are unable to attribute observed changes to the program, they will have difficulty in determining program quality.

---

* For additional information about cost-analysis approaches, see Annotated Bibliography Nos. 1, 28, and 29.

Generalizing program effects. A related issue is the extent to which the findings of an evaluation study can be generalized to other situations. The issue here is whether the program would be expected to produce similar results with, for example, a different group of participants, slight variations in the program, or changes in program personnel. The degree to which the results of an evaluation study can be generalized elsewhere is technically described as the study's *external validity*.

If evaluations are generalizable, they can provide useful information to (a) program personnel regarding the range of conditions under which the program is effective and (b) other health educators who may wish to adopt an already "evaluated" health education program. A physical fitness promotion program that works well in one setting may provide helpful guidelines for those wishing to operate other fitness programs. Typically, however, a local evaluation should be conducted once the program has been adopted.

It is important to distinguish between a program's causative power and the program's generalizability, because different information may be required to establish each factor. Procedures that limit the number of extraneous variables in the evaluation (e.g., including only males) increase internal validity but, at the same time, limit generalizability. Evaluators must try to balance the problems associated with threats to internal and external validity by selecting a data-gathering design that best addresses the information needs of program personnel as well as of those external to the program who may be interested in adopting the program elsewhere.*

Rights of Participants

Health education programs are designed to improve individuals' health and well-being. When such programs are evaluated, therefore, the focus is typically on a program's impact on human beings. Some evaluators, however, become so caught up with the importance of appraising a health education program that they overlook the rights of the individuals who take part in the evaluation. Two important rights are those of informed consent and confidentiality.

Informed consent. Evaluators, just as researchers, should be guided by a profound respect for human dignity. Therefore, they should not engage in evaluative activities that in any way demean participants. Prominent among the considerations that should guide evaluators is the concept of *informed consent*. Informed consent requires that an evaluator secure, in advance of the study, permission from the participants in an investigation to gather data from them. This consent is obtained *after* the potential participants have learned about the nature of the investigation and what their role would be, because that information may influence their decision to participate. Informed consent eliminates the possibility of making individuals unknowingly serve as subjects in an evaluation.

---

* For additional information about internal and external validity issues, see Annotated Bibliography Nos. 8, 11, 12, and 16.

Two different approaches to securing informed consent have been employed by program evaluators. The first of these, *active informed consent,* obliges an evaluator to obtain, in writing, a statement from each participant indicating that the individual is willing to participate in the evaluation. The significant aspects of the evaluation must be described in the written permission form so that potential participants are fully informed when they give their consent.

An evaluator using the second approach, *passive informed consent*, supplies descriptions of the evaluation's essentials to all program participants and provides them an opportunity to register, in writing, their unwillingness to participate in the study. In other words, when a passive informed consent approach is used, participants return the forms supplied to them only if they are *not* willing to participate in the evaluation study. Of the two approaches, the active informed consent strategy typically results in fewer participants because those individuals who do not provide consent forms must be excluded from the study. Because evaluators who conduct studies involving school-age children are obliged to secure informed consent from underage participants' parents or guardians, a passive informed consent strategy is often adopted due to the difficulty of securing active informed consent from individuals who are not participating in the program themselves.

Procedures for developing forms for both of these approaches to securing informed consent are described in Appendix B. The actual forms to be used in an evaluation would need to be more specifically relevant to the program involved.

Confidentiality. Another consideration when dealing with human subjects is the *confidentiality* of all information gathered during an evaluation. Because the evaluator is not concerned with an appraisal of individual participants but, rather, with gauging the effectiveness of a health education program, ensuring participant confidentiality usually poses no problem. Evaluators must, however, devise protective safeguards, such as anonymous completion of forms and careful handling of data, to ensure both the appearance and reality of confidentiality.*

Selecting Appropriate Measures

Although there are various approaches to program evaluation, almost all share one common feature, namely, the systematic gathering of evidence regarding a program's effects. To secure evidence of program effects, evaluators usually employ measurement instruments. Some instruments, however, are far more suitable for assessing a program's effects than others.

Criterion-referenced measurement. For more than two decades, educational measurement specialists have directed increasing attention toward an emerging form of assessment known as criterion-referenced measurement. In comparison to norm-referenced measurement, which attempts to ascertain an examinee's status in relation to the status of other examinees, criterion-referenced measurement attempts to ascertain an examinee's status in relation to a clearly defined set of behaviors. The essence of a criterion-referenced instrument is the clarity with which its accompanying descriptive materials explain what is being measured. Because norm-referenced instruments emphasize *relative* comparisons among examinees, they often do not provide a clear description of exactly what it is they are assessing. In contrast, criterion-referenced instruments are *absolute* measures, designed to determine exactly what it is that examinees can or cannot do, without reference to the performance of other examinees. Thus, criterion-referenced tests provide a clearer description of what they are measuring.

---

* For additional information about the rights of human subjects and the ethics of evaluation, see Annotated Bibliography Nos. 2, 26, and 38.

It is the clarity regarding what is being assessed that renders criterion-referenced measures ideal for the evaluation of health education programs. Consistent with the mission of providing useful information for decision makers, criterion-referenced instruments describe the precise nature of what is being measured. Hence, when criterion-referenced measures are used to gather evidence in program evaluations, decision makers can accurately interpret the evidence being supplied.*

Attributes of well-constructed measures. All instruments, whether norm-referenced or criterion-referenced, should measure what they are measuring with consistency. The consistency with which an instrument measures is known as its *reliability***. There are several different indices that can be computed to reflect an instrument's reliability. The kind of reliability data needed to appraise a measure for possible use in an evaluation study should be consonant with the way the measure will be used in that study. If a measure is to be used on a test-retest basis, for example, then information about that type of reliability is germane. If alternate forms of a test are to be used, for instance, in a pretest-posttest situation, then evidence should be available regarding alternate-forms reliability so that the evaluator can determine whether or not the two different forms are sufficiently equivalent.

It should be noted that when a health education program is being evaluated, attention should be directed to the impact of the program on a *group* of participants. Thus, the consistency to be sought when measurement instruments are used for program evaluation is consistency for a group of participants' scores. When dealing with individual participants, the measures must yield *individual* or diagnostic consistency.

A second critical attribute of a properly constructed measure is that it yields scores from which valid inferences can be drawn. An instrument is often said to be valid "if it measures what it purports to measure." Such a statement, however, is technically in error. Tests themselves are never valid or invalid. Rather, it is the *interpretations* made from test scores that are valid or invalid.

---

* For additional information about the nature and development of criterion-referenced measures, see Annotated Bibliography Nos. 7, 24, and 34.

** For information about determining the reliability of measuring instruments, see Annotated Bibliography Nos. 3, 18, 19, 23, 27, and 34.

There are several types of validity evidence, each yielding somewhat different but conceptually related indications about our ability to make valid inferences from a measure. Evidence of validity is, in the opinion of most measurement specialists, the most important consideration in judging the adequacy of measurement instruments. Program evaluators should make sure they are knowledgeable about methods of securing validity evidence.*

A final consideration in appraising the quality of measures used for program evaluation deals with the presence of *bias* in the assessment devices. During the past decade, measurement specialists have become particularly aware that many educational assessment devices contain items biased against particular subgroups, such as ethnic minorities or women. An example of a biased test item would be a knowledge question that, because of peculiarities in its content or wording, is more difficult for women to understand and answer correctly than it is for men, even though the men and women have an equivalent amount of knowledge regarding the particular concept being tested.

Another type of bias that can adversely influence examinee performance arises when test items are offensive to particular groups of individuals. For example, if a test item includes content that is seen to be derisive to members of particular ethnic groups, then examinees from those groups are not apt to perform at their best on the item. Their warranted agitation over the offensive content is likely to interfere with their responses to that item as well as to subsequent items. There are now available both judgmental and empirical techniques for detecting the presence of biased items. These approaches should be used to identify, then eradicate, bias in a measure's items.**

Finally, it is important to note that any given instrument may not possess all of the qualities discussed above. Often evaluators must choose among measures that embody some but not all of the elements described here, that is, (a) descriptive clarity, (b) reliability, (c) validity, and (d) absence of bias. Another important point is that merely because a measure is *labeled* in a particular way, for example, as criterion-referenced or as nonbiased, that does not automatically indicate that it is of sufficient quality to be used in evaluating a health education program. Scrutiny of all aspects of the measure's quality is requisite.

When to Administer Measures

Decisions regarding when to administer measures depend on the data-gathering design selected. Conceivably, there are four temporal periods during which it may be useful to obtain evaluative information about participants of health education programs. There may also be reasons for repeated measurement during some of these periods. These periods are depicted in Figure 2.1.

Pretests. Often it is useful to have information about participants prior to their starting the program. Such information, typically referred to as pretest data, may be used to identify participant needs so that instruction can be targeted directly at those areas. In addition,

---

* For information about obtaining validity evidence regarding measuring instruments, see Annotated Bibliography Nos. 3, 18, 19, 23, 27, and 34.

** For information about methods for avoiding test bias, see Annotated Bibliography Nos. 6 and 33.

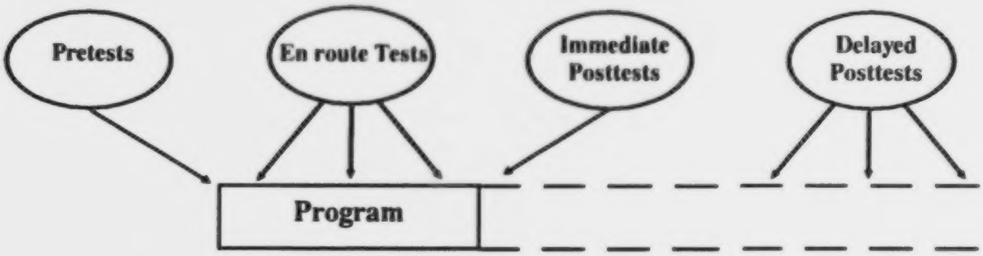


Figure 2.1: Possible measurement times in program evaluation studies

pretest data can be compared with data collected at the end of a program. Such a comparison can provide a measure of program impact.

En route tests. Measures can also be administered during a program to secure current readings on the status of participants. For purposes of formative evaluation, en route data can be used to redirect resources during the program by providing program personnel with ongoing status-checks on participants' progress. Thus, en route tests may be even more useful than tests administered at the end of the program, because en route measurement provides information while there is still time for program personnel to act on it. This type of assessment is most appropriate for programs of long duration (e.g., several months or more).

Immediate posttests. Measures are commonly administered following a program. The data from posttests can be compared with pretest data to examine changes in participants from the beginning to the end of the program. Participants' posttest performance can also be contrasted with posttest scores from participants in other programs. In addition, posttest data provide an indication of the absolute status of participants on the variables of interest at the completion of the program.

Delayed posttests. Data from delayed or follow-up posttests are often as important or more important than immediate posttest data in evaluating a health education program. Delayed posttest data might be secured, for example, several months after a program's conclusion. Far too frequently data collection efforts are limited to those times when measurement is most convenient. Ultimately, however, health educators should be interested in effecting long-term, rather than short-term, behavioral, affective, and cognitive changes. It is nearly impossible to infer such long-term changes on the basis of information gathered solely at the end of a program. As indicated in Chapter One, many of the desired changes in participants of physical fitness programs represent long-term rather than short-term objectives. For most health education programs, some follow-up measurement is usually warranted.

Clearly, it is not sensible to administer all measures at all time periods. Evaluators, in collaboration with program personnel and other interested parties, need to select a measurement scheme that focuses on the most appropriate times for gathering data. Just as it is desirable to avoid administering an excessive number of different measures, it is also necessary to avoid an excessive number of administrations. It may be useful to administer certain measures (for example, a brief behavioral self-report measure) on a continuing basis; other more time-consuming measures might be administered less frequently. Decisions about when to administer measures should be guided by common sense, attentiveness to participants' feelings, the efficient use of resources, and any conventional expectations, such as when a delayed posttest is ordinarily given.

Data-Gathering Design Options

It is sometimes thought that program evaluations must include complicated and elaborate data-gathering designs in order to yield decisive and compelling data. This is simply not the case. Program personnel and evaluators should try to conduct evaluation studies and gather data in such a way that the ambiguity of results can be reduced to a minimum. That is, evaluations must attempt to determine whether a program works and what makes it work or what prevents it from working. Data-gathering designs serve as the means to this end by setting forth the procedures to be used in exploring the nature and impact of a program.

The data-gathering design that an evaluator chooses for an evaluation will determine the inferences the evaluator can make about a program's overall impact on participants and the effectiveness of its various components. To select the best designs for evaluation studies, evaluators must have a broad knowledge of the available data-gathering design alternatives and the strengths and weaknesses associated with each. Evaluators must also work closely with program staff to determine what decisions are at issue regarding the program. No evaluation study will be perfect; every evaluation leaves some questions unanswered. Evaluators need to be clear regarding what they have learned about a program and the degree of certainty associated with their findings, and they must convey this information to appropriate audiences.

An important concept related to data-gathering designs is randomization. Randomized selection and assignment are described below, followed by brief descriptions of the most common data-gathering designs available for evaluators of health education programs.

Randomization. One technique that can prove useful to evaluators is *randomization,* which involves the selection or assignment of participants in a nonsystematic manner, such as by using a table of random numbers (found in most statistics texts). A prominent application of randomization in program evaluation is *randomized selection* of subjects. This sort of randomization is particularly important when the evaluator wishes to generalize from the results of a study to a larger population. When the participants taking part in the program to be evaluated have been selected at random from a larger population of potential participants, then the evaluator can be reasonably confident that those involved in the evaluation will be representative of that larger population. There is less likelihood that the participants being studied in the evaluation are atypical, which would make it inappropriate to generalize the evaluation's results to the population at large. Randomized selection of subjects may also be useful when there are more applicants than vacancies for a program.

Another use of randomization is to assign participants to different "treatments" or programs. If an evaluator wishes to compare the effects of different treatments, then the evaluator wants the participants in each treatment to be as equivalent as possible. To this

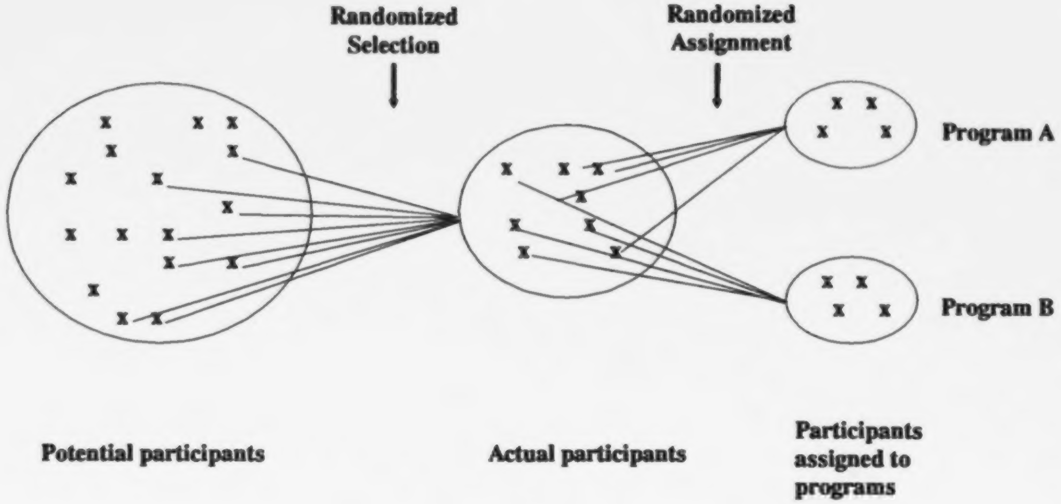


Figure 2.2: Randomized selection of participants from pool of potential participants and randomized assignment of participants to programs

end, evaluators can employ a *randomized assignment* procedure whereby individuals are randomly placed in the treatments or programs to be compared.

The two procedures of randomized selection and randomized assignment are illustrated in Figure 2.2. Note that participants are randomly selected from the pool of potential participants, and then randomly assigned to either Program A or Program B.

The use of randomization techniques does not necessarily create equivalent groups. For example, if an evaluator were to randomly assign 50 potential participants in a company's physical fitness program to treatment and no-treatment groups, it is still possible that one of the groups would contain individuals who, when pretested, were significantly different in some important aspect from those in the other group. In such instances, evaluators must rely on statistical procedures in an effort to compensate for such disparities. In most cases, however, use of randomization will create groups of sufficient equivalence that such statistical adjustments are not needed.

In practice, program personnel often may not have the luxury of constituting groups via randomized selection or assignment. For example, local school board policies might require that *all* youngsters be provided with any program regarded as potentially beneficial. When randomization is not used, it is especially important to collect and examine descriptive data about participants to determine where pre-program group differences occur and to consider the ways in which such differences may influence post-program data. Even if randomization is impossible, attempts to constitute comparison groups with individuals as equivalent as possible can help minimize the influence of preexisting participant differences.*

Seven different data-gathering designs of potential utility for evaluators of health education programs will be presented below. Each data-gathering design will be described and depicted schematically. Some of the major factors involved in the selection of data-gathering designs will be addressed.

The case-study design. Consider a six-month health education program aimed at modifying participants' knowledge about the effects of physical fitness on health. If participants' knowledge were measured only at the close of the program, we could describe the data-gathering approach as a *case-study design* and represent it schematically as shown in Figure 2.3.

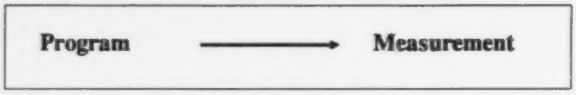


Figure 2.3: Case-Study Design

If this were the design employed in an evaluation, what could an evaluator tell about the program's impact on participants' knowledge? How confident would an evaluator be that participants' knowledge about the effects of physical fitness was attributable to the program?

It would be difficult, with confidence, to attribute any effects to the health education program. The program, indeed, may have been totally ineffectual. In fact, participants' post-program knowledge might be identical to their knowledge before the program. The participants could be demonstrating knowledge that they brought to the program, not that they acquired during the program. Because we have no measure of participant knowledge prior to the program, we cannot distinguish between preexisting knowledge and knowledge acquired as a result of the program. Hence, with the case-study design, it may be impossible to determine whether the program had any impact on participants.

Even though attributions of causality are often unwarranted, it may be possible to secure useful program evaluation data with such a data-gathering design. Suppose, for example, that a health education program is promoting a body of knowledge so advanced that few, if any, individuals would be familiar with it. In such a setting, one could assume that participants' post-program knowledge is attributable to the program's impact because participants would almost certainly not have acquired the knowledge without the program. It might not be worth the resources necessary to implement a data-gathering design capable of conclusively demonstrating that participants began the program unfamiliar with the knowledge being promoted.

---

* For additional information about randomization, see Annotated Bibliography Nos. 8 and 25.

This example illustrates an important data-gathering consideration, namely, that the chief mission of data-gathering designs is to *rule out plausible rival explanations,* that is, explanations other than the program's impact that might account for the post-program status of participants. If there is reason to believe that participants' pre-program status may account for their post-program status, then a data-gathering design should be selected that permits the evaluator to rule out this rival explanation.

The one-group pretest-posttest design. Now suppose that, to avoid the major shortcoming of the case-study design, an evaluator measures participants' behavior both before and after a health education program. This data-gathering approach can be described as a *one-group pretest-posttest design* and can be represented as shown in Figure 2.4.

Measurement → **Program** → **Measurement**

Figure 2.4: One-Group Pretest-Posttest Design

Assume an evaluator uses the one-group pretest-posttest design and that the data reveal a substantial shift toward more desirable behaviors between the initial and the final measurement. Can this change in behaviors be ascribed to the program? Unfortunately, the evaluator cannot be sure. There are many other factors, totally unrelated to the program, that may have influenced participants' behaviors. For instance, if a physical fitness program emphasized the relationship between exercise and health, and at the same time a new study confirming the idea that exercise decreases the likelihood of illness received attention in the national news, such an event may have influenced participants' views regarding exercise and health. Evaluators of programs that serve children must also consider the possible effects of maturation during the time the program is offered. Participants' increased maturity may cause pre-program to post-program shifts in behaviors. The program itself may have contributed nothing to the measured shift of behaviors. Such extraneous factors decrease the evaluator's ability to draw defensible conclusions about the program's impact.

As was true with the case-study design, however, if there are no plausible rival explanations for the posttest results, the one-group pretest-posttest design can be suitable for the task at hand. In fact, this simple yet serviceable design is often used in formative evaluation.

The one-group pretest-posttest design requires measurement before as well as after a program. This points to a commonly accepted but often overlooked principle of effective program evaluation. Evaluation is most effective when it is initiated at the beginning of a program. If evaluators are not called in until the end of a program, they may be hampered in their efforts to design a credible program evaluation.

The nonequivalent control/comparison group design. Program evaluators can eliminate some of the more common rival explanations for changes in participants' behaviors by using data-gathering designs in which either comparison or control groups are employed. The use of a control group (untreated individuals) or a comparison group (individuals receiving a different program) requires two groups that are assumed to be relatively similar (before the program) on all related variables. When using these designs, the evaluator should attempt to secure two groups that are as similar as possible. Because the two groups are not randomly assigned to the two conditions, however, they cannot be assumed to be *equivalent*, hence the design's designation as a "nonequivalent" control or comparison group design.

In the control-group version of this design, only one of the groups is given the program to be evaluated; the other group is left untreated. This data-gathering design, known as the *nonequivalent control group design,* is illustrated in Figure 2.5.

Group 1: **Measurement** → **Program** → **Measurement**
Group 2: **Measurement** → **Measurement**

Figure 2.5: Nonequivalent Control Group Design

In this design, a control group (Group 2) is assessed before and after the program, but it never receives the program itself. Assuming that the groups were similar before the program, if the program participants' behaviors change while the behaviors of those in the control group remain the same, the evaluator can be reasonably confident that the program caused the change.

The use of an *untreated* control group may strike some health educators as a particularly unsavory data-gathering ploy. After all, health educators design their programs to benefit participants. To withhold such programs from individuals, even for the important purpose of evaluating the program's effectiveness, seems downright reprehensible. Yet, the individuals from whom the program is withheld, that is, the members of the control group, can be given the program *subsequently,* as soon as the evaluation study has been concluded. Also, in some situations there are more program applicants than can be accommodated, and, therefore, some prospective participants must be denied access to this program under any circumstances. Those who are not admitted to the program could be used as a control group, and admitted to the program the next time it is offered.

A variation of the nonequivalent control group design involves the use of a comparison group, that is, a group receiving a different program or a different treatment. Program evaluators frequently find themselves studying the quality of two or more competing programs. Thus, the evaluator focuses on the relative virtues of two or more different programs rather than on a contrast between a single program and an untreated control group. A schematic depiction of a *nonequivalent comparison group design,* in this instance contrasting two different programs, is presented in Figure 2.6. As indicated above, more than two groups can be employed when using a nonequivalent comparison group design. An evaluator using this design can be fairly certain that, if the groups were similar before the program, any differences in post-program behaviors are due to the differential impact of the two programs.

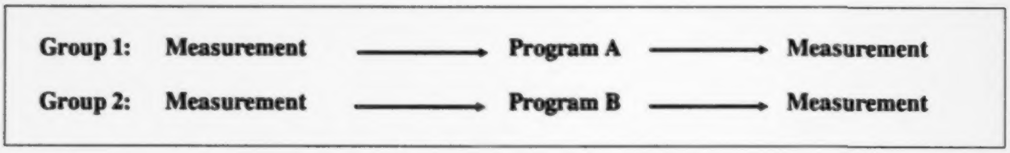


Figure 2.6: Nonequivalent Comparison Group Design

There are, however, potential problems with the nonequivalent control/comparison group designs. It may be that the initial measurement was *reactive*. A reactive measurement is one that, by itself or in combination with the program, influences participants' behavior. Attitude inventories and self-report questionnaires about behavioral practices are notoriously reactive. For example, a questionnaire administered before the program might alert participants to the importance of a desired behavior. This would heighten their attentiveness when the program dealt with content related to that behavior and, as a consequence, influence their performance on the second measurement.

Moreover, measurement is expensive. Measuring the status of control groups requires valuable evaluation resources. Time and money can often be better spent studying the program being evaluated rather than studying a no-treatment control group of little intrinsic interest. Health educators should not ritualistically employ control groups in their designs if the questions at issue can be answered without the use of untreated groups.

The pretest-posttest control/comparison group design. There are two data-gathering designs that are of particular value to program evaluators if randomized assignment is possible. The first of these is the *pretest-posttest control group design*, illustrated in Figure 2.7.

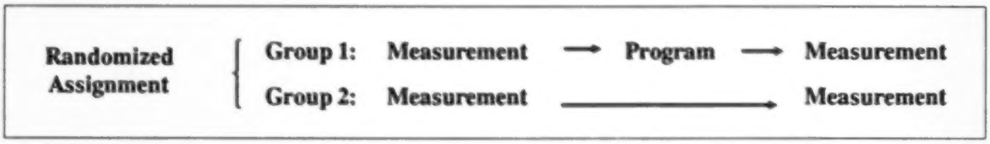


Figure 2.7: Pretest-Posttest Control Group Design

The difference between this design and the previously considered nonequivalent control group design is, of course, the randomized assignment of subjects to the two groups. This feature of the design is a particularly important one, because creation of two or more groups using randomized assignment is an effective way of promoting equivalence between the groups, especially if the number of subjects in each group is large (say, 30 or more). Equivalence of groups at the beginning of the program strengthens the inference that any differences at the conclusion of the program are due to program impact.

By using comparison groups, that is, two or more program groups, instead of an untreated control group, the evaluator would be using a *pretest-posttest comparison group design*, shown in Figure 2.8.

Randomized Assignment
Group 1: Measurement → Program A → Measurement
Group 2: Measurement → Program B → Measurement

Figure 2.8: Pretest-Posttest Comparison Group Design

Because pretests are used in both of these designs, the possibility of reactive pre-program measures is still present. For situations in which reactivity is of great concern, a different data-gathering design, described next, has much appeal.

The posttest-only control group design. In situations where a measure is likely to be reactive, the evaluator can rely on a clever data-gathering design that effectively dodges the reactivity problem. This *posttest-only control group design* is depicted in Figure 2.9. This design is the same as the pretest-posttest control group design, except that there is no pretest.

Randomized Assignment
Group 1: Program → Measurement
Group 2: Measurement

Figure 2.9: Posttest-Only Control Group Design

In this design, neither Group 1 nor Group 2 is pretested, but because of random assignment the groups can be considered equivalent prior to Group 1 receiving the program. Not pretesting Group 1 effectively avoids a pretest's potentially reactive effect on program participants. To assess the impact of the program, it is possible to contrast the *posttest* performances of Groups 1 and 2. As with the other control group designs, the untreated control group could be given the program the next time it is offered.

The basic dividend of the posttest-only control group design is that by measuring an untreated, randomly assigned control group, the evaluator secures an estimate of how program participants would have responded on a pretest, but without introducing the potentially reactive effects of a pretest. Although the diagram for this design suggests that the measurements be made for both groups at the conclusion of the program, it is possible to measure the untreated control group earlier if that seems advisable.

Multiple measures over time. There are certain situations in which health educators may wish to appraise the effects of their programs on the basis of periodic measurements, for example, by using regularly administered questionnaires or data that are routinely recorded. For instance, suppose when evaluating an employee injury prevention program, the evaluator was interested in the number of job-related injuries occurring in a company. Assuming that such information is available from the firm's health records, the evaluator might study records at periodic intervals before, during, and after the program. By observing the frequency of job-related injuries during different time intervals, the evaluator would have valuable information regarding program effects.

A number of the most commonly used data-gathering designs have been described. There are other, more complex designs than those treated here.* Complexity, however, is rarely an asset if a more straightforward design is appropriate.

Sampling Considerations for Data Collection

The data-gathering requirements of an evaluation can become a burdensome intrusion into an ongoing health education program. Accordingly, evaluators should conduct their data-gathering activities in the least intrusive manner possible. One way to minimize an evaluation's intrusiveness is by relying on sampling techniques, such as person-sampling and item-sampling, each of which is described below.

Person-sampling. To estimate how a large group of people would respond on a particular measure, it is not necessary to administer the measure to all the individuals in the group. Instead, a smaller group can be selected. This smaller group can be either a *simple random sample* or a *stratified random sample,* that is, a sample stratified on the basis of program-relevant factors such as age, sex, and socioeconomic status. Assuming that the sample is randomly selected, the evaluator can estimate the status of the total group based on the responses of the sample.

Suppose, for example, that the evaluator wants to use a measure to determine participants' knowledge about planning a safe exercise program. Assuming that there is a reasonably large number of program participants, say 50 or so, the evaluator could randomly select half of the participants and administer the measure to this group only. In essence, this approach allows the evaluator to infer how the total group of participants would score on the measure, even though only half of the participants completed it. Thus, it is possible to estimate total group performance with only half the amount of participant time required for data gathering.

Using a similar sampling procedure, evaluators can administer two or more measures at once in the time it takes to administer one. Suppose that two measures are to be given to program participants. The evaluator can randomly assign one measure to half of the participants and the other measure to the remaining participants. Each participant needs to respond to only one measure, but the evaluator can derive defensible estimates of how all the participants would have responded on both instruments.

---

* For additional information about evaluation design options, see Annotated Bibliography Nos. 8, 11, 22, 23, and 35.

Item-sampling. In addition to sampling persons, as in the previous examples, it is also possible to sample items, so that different sets of items from a program evaluation measure are randomly selected to be administered to different persons. Using this approach, the evaluator gives each participant only a sample of the items on any particular measure. For example, suppose a program evaluator wishes to administer a 30-item test. Given 60 participants in the program, the evaluator could divide the test into three sets of 10 items each and administer each set of 10 items to 20 different participants. In this way, the total group's performance on the whole test can be estimated. This approach to data gathering requires only one-third of the time that would have been required to administer the total 30-item test to all participants.

Sample size. Given the relatively small number of participants in some health education programs, is it really appropriate to sample either persons or items? How large must groups be before these sampling procedures can be sensibly used? Unequivocal answers to these questions do not exist. Some texts on sampling provide rules of thumb for estimating the size of samples needed for detecting group differences in relation to the magnitude of differences sought and the nature of the groups being sampled. At best, though, these rules provide only rough estimates. It is important to recognize that the task of identifying a sufficiently large sample is more difficult than usually thought.

The variability of participants' anticipated performance on the measures is the primary determiner of the sample size necessary. If it is expected that participants' scores on a test will be relatively homogeneous, a smaller number of respondents will be needed than if participants' scores are expected to vary widely. Thus, if on a measure of knowledge about the prevention and care of injuries, for example, some of the participants are expected to know many techniques and others are expected to know very few, reasonably large numbers of participants (e.g., 20) should respond to any one item.

Intuitively, one recognizes that when working with a very small group of program participants, the use of these sampling techniques is risky. For instance, if there were only 15 participants in a program, few evaluators would try to split these participants into three groups of five each for purposes of taking different sets of items. Even though each group represents one-third of the total population, there is too much likelihood that a sample of five individuals would not properly represent the total group. One or two atypical participants in a five-person group would render the group's average performance unrepresentative of how the larger group would have performed.

It should be noted that when employing procedures such as person-sampling or item-sampling, an evaluator is focusing on a group of participants *in the aggregate*. Because evaluations are typically concerned with the effects of programs on groups of participants, the use of sampling procedures is usually appropriate. If, however, program personnel need individual data on all examinees, then sampling should obviously not be employed.*

---

* For additional information about sampling procedures, see Annotated Bibliography Nos. 9 and 10.

Data Analysis

A frequent question asked of an evaluator is whether a study's results are statistically significant. For example, could the observed changes in program participants' knowledge or behavior from pretest to posttest have occurred simply by chance? Statistical tests are used to answer this type of question. Consideration of statistical analysis procedures, however, is beyond the scope of this handbook. Thus, just a few comments will be made here regarding data analysis. Because there are many subtle choice-points in the statistical analysis of evaluation data, evaluators who are not well versed in at least the more common statistical procedures should probably enlist the aid of someone who is.

There are two basic classes of statistics, namely, descriptive statistics, such as the mean, and inferential statistics, such as the *t* test. *Descriptive statistics* help evaluators portray a group's performance on a given measure. For example, an evaluator might describe a set of participants' scores via the mean score (the scores' central tendency) and standard deviation of the scores (the scores' variability). Because the mean and standard deviation are frequently used, program evaluators should know how to calculate and interpret them. Any introductory statistics book for the social sciences will serve as a reference for this information. *Inferential statistics* help evaluators determine whether an observed difference between pre-program and post-program scores is *statistically significant,* that is, whether such a difference could have occurred because of chance alone. If the probability is small that the results are due to chance, the evaluator can, with reasonable confidence, attribute the results to the program.

Statistical significance, however, does not imply *practical significance*. A small difference between the average scores of two groups can be statistically significant, particularly when large numbers of participants are involved, yet be of no practical consequence whatsoever. Health educators will need to make sensible determinations regarding whether the magnitude of an observed difference, even though statistically significant, is sufficiently important to warrant action. In other words, although evaluators of health education programs should often carry out statistical significance tests, they should not be unduly swayed by the results of such analyses. Common sense must always be applied in interpreting the meaning of a statistically significant result.*

Reporting Results

Reporting the results of an evaluation study is a more difficult undertaking than is usually recognized. Considerable attention must be given to the procedures employed to report the results of health education program evaluations. When reporting evaluation results, as when focusing and planning the evaluation, the evaluator must be responsive to the needs of program decision makers. A few key considerations should be kept in mind when reporting evaluation results.

Evaluators must report their results to decision makers in a timely fashion. It does no good to deliver an evaluation report several weeks after key program decisions had to be made. Evaluators must also be careful to disseminate their findings to all appropriate audiences. If possible, an evaluator should circulate the preliminary draft of a program evaluation report to program personnel so that they can react to its accuracy and objectivity.

---

* For additional information about data analysis, see Annotated Bibliography Nos. 25, 36, 39, 43, and 45.

The decision makers whom evaluators are assisting may have scant experience with quantitative data. As a consequence, complicated statistical presentations may be of little value to them. Evaluators should select data-presentation procedures that will match the technical sophistication of the decision makers involved. In any evaluation report, there is nothing wrong with simple graphs or "percentage correct" tables. The more intuitively comprehensible the data-presentation techniques, the better they are. Program evaluators should provide straightforward presentations of data without fearing that such approaches will be regarded as too elementary. Adequate technical back-up can be appended as necessary to the final report.

Evaluators should not be reluctant to make speculations based upon their knowledge about a program, but these conjectures should be identified as such. Similarly, if any of the evaluation's findings are equivocal, the evaluator should inform concerned audiences of this fact. Honesty and objectivity are the hallmarks of effective evaluation reporting.

In addition, because decision makers are typically busy people, evaluators should strive for reasonable brevity in their reports. The preparation of executive summaries to accompany lengthy reports is a useful practice. Voluminous evaluation reports are almost certainly destined to go unread. Terse, easily read reports are much more likely to make an impact on decision makers.

The whole thrust of the evaluation enterprise is to facilitate better decisions. Decision making will *not* be illuminated by complex, lengthy, or otherwise incomprehensible presentations of evaluation results. The quality of decision making can be enhanced only if an evaluation's results are reported in a way that can be clearly understood.*

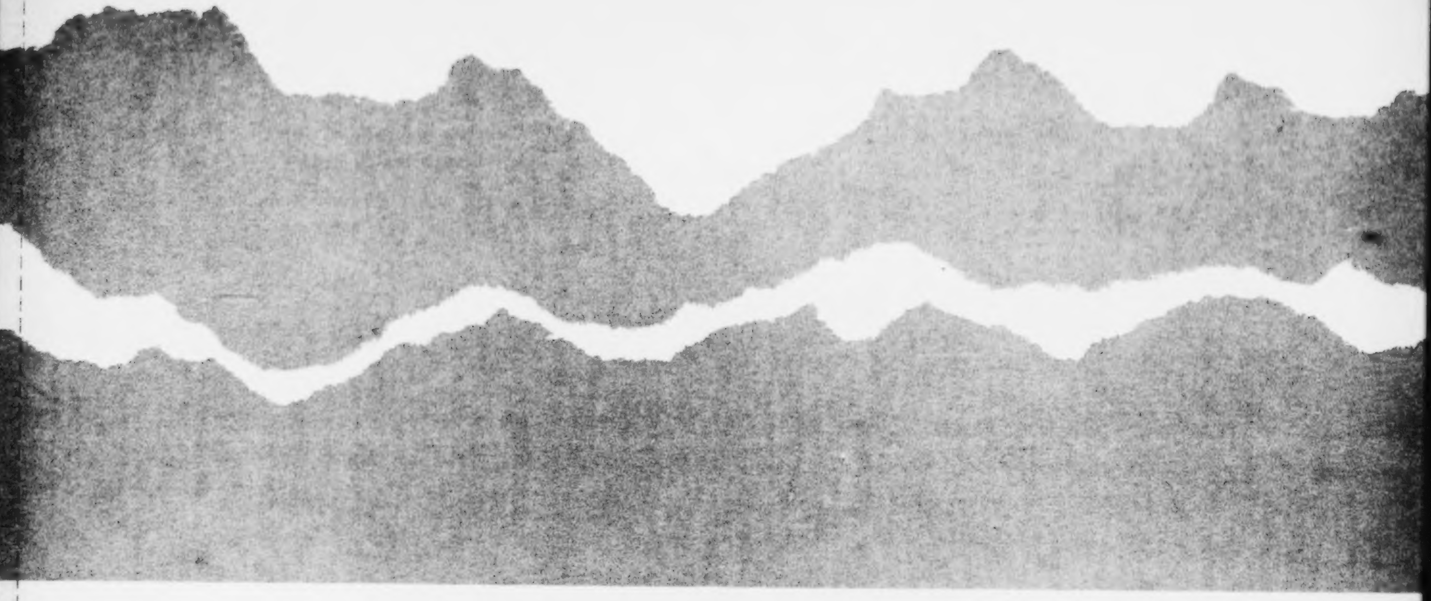
Reprise

In this chapter, a number of issues almost certain to be encountered by evaluators of physical fitness programs were considered. Because this handbook supplies a number of measures to be used in the evaluation process, special attention was given to the role of such measures in program evaluation. Evaluators desiring more detailed treatments of the topics covered in this chapter will find appropriate sources in the Annotated Bibliography.**

---

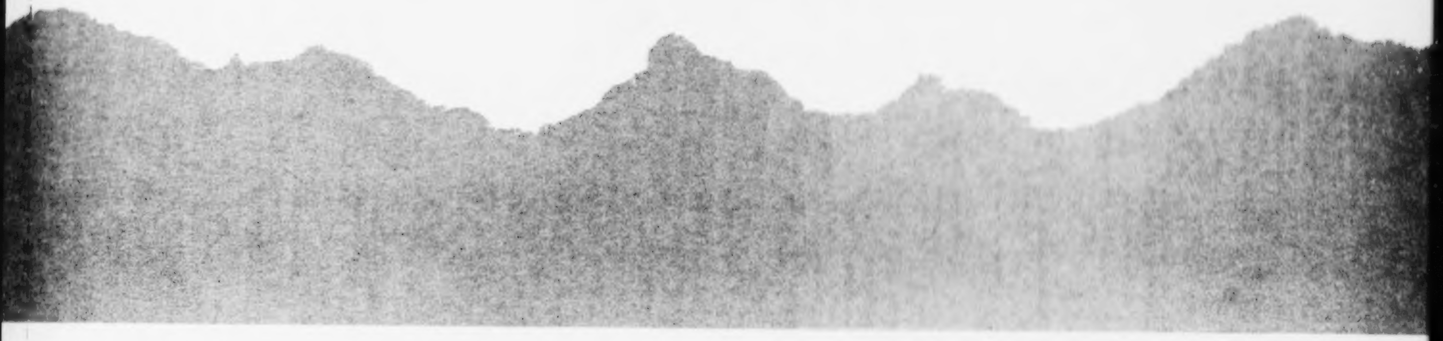
* For additional information about reporting the results of an evaluation, see Annotated Bibliography Nos. 5, 23, 26, and 35.

** For additional information about program evaluation, see Annotated Bibliography Nos. 5, 13, 16, 20, 23, 32, 41, 46, 49, and 51.



CHAPTER THREE

Physical Fitness Promotion Measures



Overview of Measures

Category	Title	Target Group	Description	Page No.
Behavior	**Weekly Activities Index**	Adults Adolescents	Assesses frequency and intensity of exercise.	37
	Injury Checklist	Adults Adolescents	Assesses location and severity of exercise-related injuries in the past month.	43
Knowledge*	**Facts About Exercise**	Adults	Assesses knowledge of the effects of exercise on the body.	46
	Exercise Facts	Adolescents Preadolescents		52
	Planning a Safe Exercise Program	Adults	Assesses knowledge about planning a safe exercise program.	56
	Planning to Exercise Safely	Adolescents Preadolescents		62
Skill	**Selecting an Exercise Program**	Adults	Assesses ability to select an appropriate exercise program.	66
	Preventing and Caring for Injuries	Adults	Assesses ability to prevent and care for exercise-related injuries.	80
	Exercising Safely	Adolescents *Preadolescents*		90

* The information eligible for inclusion in the knowledge measures is provided in Appendix A as amplified content descriptors.

Category	Title	Target Group	Description	Page No.
Affective	**Ideas About Decisions**	Adolescents Preadolescents	Assesses belief in the value of careful decision making.	96
	Effects of Exercise	Adults	Assesses belief in the positive effects of exercise.	99
	Beliefs About Exercise	Adolescents Preadolescents		101
	Exercising Regularly	Adults	Assesses perceived ability to exercise regularly in a variety of situations.	103
	Intention to Exercise	Adults	Assesses intention to begin or maintain a regular exercise program.	106
	Attitude Toward Work	Adults	Assesses attitudes toward work.	109

WEEKLY ACTIVITIES INDEX

This behavior measure assesses participants' frequency and intensity of exercise during the past seven days. This measure is appropriate for adults and adolescents.

If this measure seems useful, you might also want to consider administering the **Intention to Exercise** measure, which is an affective measure examining participants' intention to continue exercising or begin exercising.

PURPOSE

Information about participants' frequency and intensity of exercise may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show that participants have a low level of activity, thus indicating a need for instruction in the positive effects of exercise.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' frequency and intensity of exercise.

PROCEDURES

This instrument should be administered both at the beginning and at the end of the program. If an individual's last seven days was highly atypical, then the previous week should be used as the basis for completing the index.

SCORING AND ANALYSIS

MET[*] values are assigned to activity categories (Questions 4 and 5) as follows:

Moderate = 4 METS
Hard = 6 METS
Very Hard = 10 METS

This index can be scored by multiplying the total number of *minutes* spent on activities reported within a category by the assigned MET value. The MET values from the three categories, including household and occupational activities, should be added to determine the respondent's total MET value or the energy expended for specified activities during a seven-day period.

---

[*] A "MET" is equal to the energy expended or oxygen needed at rest. Therefore, if an activity is equivalent to 4 METS then it requires 4 times the amount of energy expended at rest. According to the American Heart Association (1973), 100-240 minutes a week of activities (5 METs or above) is needed to maintain a moderate to very low risk of developing coronary heart disease. For more information regarding METS see Chapter One *Specific Physical Fitness Concerns*.

WEEKLY ACTIVITIES LIST

The following questions ask about your activities during the past *7 days*.

1. Compared to usual, how active were you during the last 7 days? (*Circle one*)

1	2	3	4	5	6	7
Much less active than usual			**About the same**			**Much more active than usual**

2. About how many times did you exercise enough to work up a sweat during the past 7 days? ____________________

3. Use the chart below to indicate whether you did any stretching or strength exercises during the past 7 days. (Start with this time, one week ago today.) If you did, indicate the *total* number of minutes you spent on each activity. (Exclude waiting time, breaks, etc.) Also indicate the part(s) of the body for which the exercise was intended.

Activities	(Circle one)	Total Number of Minutes	Body Parts (check all that apply)
Stretching	Yes No		☐ upper body ☐ trunk ☐ lower body
Strength Exercises (such as push-ups, sit-ups, Nautilus)	Yes No		☐ upper body ☐ trunk ☐ lower body

4. The next two pages ask about the *moderate*, *hard*, and *very hard* activities that you may have engaged in during the past 7 days. Do not record light activities such as bowling. For each activity you engaged in, indicate the number of different times you performed the activity and the *total* number of minutes you spent on the activity. (Exclude waiting time, breaks, etc.)

   If you participated in any exercise activities (other than light activities) that are not listed, add them to the appropriate list. To decide where to add additional activities, examine the activities in each list and consider how difficult it would be to perform the activity for a long time.

MODERATE ACTIVITIES	**Number of Times**	**Total Number of Minutes**
Baseball/Softball		
Brisk walking (15-20 minutes per mile)		
Calisthenics		
Dancing (social)		
Golf (no cart)		
Hiking		
Table tennis		
Volleyball (6-person)		
Other *moderate* activities: ________________ ________________ ________________		

HARD ACTIVITIES	**Number of Times**	**Total Number of Minutes**
Aerobic dancing		
Bicycling (leisure)		
Doubles racquet sports		
Doubles handball		
Downhill skiing		
Light jogging (13-14 minutes per mile)		
Skating (roller or ice)		
Stair climbing (for exercise)		
Swimming		
Volleyball (2-person)		
Water skiing		
Other *hard* activities: ____________________ ____________________ ____________________		

VERY HARD ACTIVITIES	**Number of Times**	**Total Number of Minutes**
Basketball		
Bicycling (racing or training)		
Cross country skiing		
Jogging or running (12 minutes per mile or faster)		
Judo/Karate		
Rope skipping		
Rowing		
Singles racquet sports		
Singles handball		
Soccer		
Swimming (laps)		
Touch football		
Other *very hard* activities: ____________________ ____________________ ____________________		

5. This question asks about additional *occupational* or *household* activities that you may have engaged in during the past 7 days. Please record the amount of time, if any, you spent on the types of activities listed. Light activities such as dusting are not included in this question. (Do *not* duplicate information provided in Question 4.)

	TOTAL NUMBER OF MINUTES	
ACTIVITIES	**At home**	**At work**
Moderate household or occupational activities (such as cleaning windows, mopping, house painting)		
Hard household or occupational activities (such as heavy gardening, scrubbing floors, construction work)		
Very hard household or occupational activities (such as digging, shoveling, carrying heavy loads)		
Bicycling or walking for transportation		
Stair Climbing (Total number of flights climbed up)		

The last two questions are about your general level of activity.

6. Compared to others your same age and sex, how active do you consider yourself?

1	**2**	**3**	**4**	**5**	**6**	**7**
Very inactive			**About the same**			**Very active**

7. Did you do any of the following activities regularly throughout the past three months?

A.	Jog or run at least 10 miles per week.	Yes	No
B.	Play strenuous racquet sports at least 5 hours per week (singles tennis, racquetball, etc.).	Yes	No
C.	Play other strenuous sports at least 5 hours per week (basketball, soccer, or other sports involving running).	Yes	No
D.	Ride a bicycle at least 50 miles per week.	Yes	No
E.	Swim at least 2 miles per week.	Yes	No

INJURY CHECKLIST

This behavior checklist examines the location and severity of exercise-related injuries participants have had during the past month. This measure is appropriate for adults and adolescents.

If this measure seems useful, you might also want to consider administering the **Preventing and Caring for Injuries** skill measure, which assesses participants' skill in preventing and caring for injuries.

PURPOSE

Information about participants' location and severity of exercise-related injuries and the activities which commonly cause injury may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show that participants have a higher than normal amount of injuries, thus indicating a need for instruction in the use of injury prevention techniques.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants frequency and severity of exercise-related injuries.

PROCEDURES

This instrument should be administered both at the beginning and the end of the program. If the program is fairly long (several months or more), the instrument may be given as it exists in this handbook. If the program is shorter than two months, it is possible that the program will not produce the behavior changes measured by this instrument. Instead, programs of shorter duration should use this measure for the needs assessment purposes.

SCORING AND ANALYSIS

- **Average Number of Injuries**

  To determine the average number of injuries experienced by all program participants, add the total number of injured body parts checked in response to the question "Have you injured this part?" for all participants. Divide this sum by the number of participants.

- **Percentage of Participants with Specific Injuries**

  To determine the percentage of participants who injured a specific body part (e.g., ankle), add the number of participants who checked ankle in response to the question "Have you injured this part?" Divide this sum by the total number of program participants and multiply by 100.

- **Cause of Injury**

  To determine which activities participants were engaged in when injured, compile a list of the different activities cited by participants for each body part.

- **Severity of Injury**

  To determine the percentage of participants who either (a) missed work/school, (b) reduced their exercise program, or (c) obtained medical care as a result of a body part injury, add the number of participants who checked at least one of the three columns for each injured body part (e.g., ankle). Divide this sum by the number of participants who indicated that they injured their ankle (in response to the question "Have you injured this part?") and multiply by 100.

Example:

Imagine that there are 10 participants in the program and 8 of them checked that they injured their ankle. Of those 8 participants, 4 had to either miss work/school, reduce their exercise program or obtain medical care. Divide 4 by 8 and multiply by 100 to find that 50% of those participants who injured their ankle did so badly enough to require extra attention.

INJURY CHECKLIST

This survey asks about exercise-related injuries that you may have had during the past month. If you have *not* had any exercise-related injuries, place a check in the box below and do not complete the rest of this checklist:

☐ **I have not had any exercise-related injuries in the past month.**

If you have had any exercise-related injuries in the past month, please complete the rest of this checklist.

Place a check (√) next to the body part(s) that you injured. For every body part you injured, indicate the activity you were engaged in when the injury occurred. Finally, write "Yes" or "No" to indicate whether the injury caused you to miss work/school, *reduce* your exercise program or obtain medical care.

Body Part	Have you injured this part?	What activity were you engaged in when the injury occurred?	Has this injury caused you to . . .		
			miss work/ school?	*reduce* your exercise program?	obtain medical care?
Foot/toes					
Ankle					
Knee					
Leg					
Hip					
Groin					
Back					
Shoulder					
Neck					
Arm					
Hand/fingers					
Wrist					
Eye					
Head					
Nose					
Other (please specify)________ ____________ ____________					

FACTS ABOUT EXERCISE

(FORMS A & B)

This knowledge measure examines what participants know about the physiological effects of exercise including the effects of exercise on body composition, weight reduction, and the cardiorespiratory, muscular, and skeletal systems. This measure is appropriate for adults.

PURPOSE

Information about participants' knowledge of the effects of exercise may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show limited knowledge regarding the effects of exercise on the body, thus indicating a need for instruction in that area.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' knowledge of the effects of exercise.

PROCEDURES

Because the equidifficulty of the forms has not been established, it is better not to give all participants Form A as a pretest and Form B as a posttest. Instead, choose either of the following methods.

- Review Forms A and B and select one. Give all participants the selected form both before and after the program. Alternatively, select 20 items from the two forms and construct a measure most consistent with your program emphasis. Then administer the "new" form both before and after the program.
- Give Form A to half of the incoming participants and Form B to the remaining half. To distribute the forms randomly, order them "ABABAB" and hand them out. Following the program, give each participant the form not previously taken. For example, if a participant was given Form B before the program, then that participant should be given Form A following the program. This approach eliminates the possibility that examinees will be sensitized to the specific facts to be learned from the program.

SCORING AND ANALYSIS

The answer keys for the two forms are provided below:

Item No.	Form A	Form B
1	F	F
2	T	F
3	T	T
4	T	T
5	T	F
6	F	F
7	F	F
8	T	T
9	F	T
10	F	F
11	F	T
12	T	F
13	F	F
14	F	T
15	F	T
16	T	T
17	T	F
18	T	T
19	F	F
20	T	F

The measures should be scored by counting the number of correct answers for each participant. Items marked "Don't Know" or left blank should be scored as incorrect. Count the number of correct answers for each participant. Next, total the correct answers for the group and divide by the number of participants in the group. The mean number of correct answers and the standard deviation can be used to summarize participant performance on the measure. Means and standard deviations from before and after the program can be compared to determine changes in participants' knowledge.

FACTS ABOUT EXERCISE

Form A

This test has 20 statements about the effects of exercise. Put a check to show whether you think each statement is TRUE or FALSE. If you don't know whether a statement is true or false, put a check under DON'T KNOW.

True	False	Don't Know		
( )	( )	( )	1.	Strength exercises are a good way to improve your heart/lung fitness.
( )	( )	( )	2.	The greatest heart/lung improvements from exercise generally occur at the beginning of a fitness program.
( )	( )	( )	3.	The benefits of regular exercise are gradually lost if exercise is not continued.
( )	( )	( )	4.	Flexibility exercises can help people avoid lower back problems.
( )	( )	( )	5.	Exercising joints in the body increases their flexibility.
( )	( )	( )	6.	Regular exercise usually increases a person's resting blood pressure.
( )	( )	( )	7.	All sports provide the same benefits.
( )	( )	( )	8.	In general, older people benefit from exercise as much as young people.
( )	( )	( )	9.	Exercising a particular area of the body is a good way to reduce the body fat in that area.
( )	( )	( )	10.	The longer and harder a person exercises, the better it is.
( )	( )	( )	11.	A person must exercise every day in order to become physically fit.
( )	( )	( )	12.	While exercising, a heavier person uses more calories than a lighter person.

True	**False**	**Don't Know**		
( )	( )	( )	13.	Regular exercise strengthens muscles, not bones.
( )	( )	( )	14.	Even a single exercise session can have a lasting effect on the heart/lung system.
( )	( )	( )	15.	People who are slim don't need to exercise to be physically fit.
( )	( )	( )	16.	Physically active people are less likely than inactive people to have a heart attack.
( )	( )	( )	17.	The body continues to use more calories for several hours after exercising has stopped.
( )	( )	( )	18.	People must increase their heart rate during exercise to improve heart/lung fitness.
( )	( )	( )	19.	Muscular strength is the ability to move a heavy weight many times.
( )	( )	( )	20.	Regular, moderate exercise is a good way to stay healthy and avoid disease.

FACTS ABOUT EXERCISE

Form B

This test has 20 statements about the effects of exercise. Put a check to show whether you think each statement is TRUE or FALSE. If you don't know whether a statement is true or false, put a check under DON'T KNOW.

True	False	Don't Know	
( )	( )	( )	1. In general, young people benefit more from regular exercise than older people.
( )	( )	( )	2. Regular exercise increases the heart rate at rest.
( )	( )	( )	3. Physically active people are less likely than inactive people to develop high blood pressure.
( )	( )	( )	4. Strength exercises for stomach muscles will help reduce the risk of lower back problems.
( )	( )	( )	5. Muscles that are not exercised turn into fat.
( )	( )	( )	6. Most people can maintain good flexibility without performing flexibility exercises.
( )	( )	( )	7. To be most beneficial, an exercise program should be as difficult as possible.
( )	( )	( )	8. Even healthy people need to exercise to be physically fit.
( )	( )	( )	9. Regular exercise helps the body use fat.
( )	( )	( )	10. It is bad to exercise while dieting because exercise increases one's appetite.
( )	( )	( )	11. Regular exercise can firm muscles.
( )	( )	( )	12. The best measure of body composition is weight.
( )	( )	( )	13. Regular exercise decreases the amount of oxygen the body can use while exercising.

Facts About Exercise (Form B), p. 2

True	**False**	**Don't Know**		
( )	( )	( )	14.	Regular exercise can help keep blood vessel linings clear.
( )	( )	( )	15.	Muscular endurance is the ability to move an object many times.
( )	( )	( )	16.	Regular exercise can slow down the natural decline in heart/lung fitness.
( )	( )	( )	17.	Considerable weight loss usually occurs during the first few weeks of an exercise program.
( )	( )	( )	18.	Regular exercise can increase the amount of blood pumped with each heartbeat.
( )	( )	( )	19.	Lifting weights to develop muscular endurance makes muscles big and bulky.
( )	( )	( )	20.	Regular exercise rarely improves blood circulation.

EXERCISE FACTS

FORMS A & B

This knowledge measure examines what participants know about the physiological effects of exercise including the effects of exercise on body composition, weight reduction, and the cardiorespiratory, muscular, and skeletal systems. This measure is appropriate for adolescents and preadolescents.

PURPOSE

Information about participants' knowledge of the effects of exercise may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show limited knowledge of the physiological effects of exercise, thus indicating a need for instruction in that area.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' knowledge of the effects of exercise.

PROCEDURES

Because the equidifficulty of the forms has not been established, it is better not to give all participants Form A as a pretest and Form B as a posttest. Instead, choose either of the following methods.

- Review Forms A and B and select one. Give all participants the selected form both before and after the program. Alternatively, select 15 items from the two forms and construct a measure most consistent with your program emphasis. Then administer the "new" form both before and after the program.
- Give Form A to half of the incoming participants and Form B to the remaining half. To distribute the forms randomly, order them "ABABAB" and hand them out. Following the program, give each participant the form not previously taken. For example, if a participant was given Form B before the program, then that participant should be given Form A following the program. This approach eliminates the possibility that examinees will be sensitized to the specific facts to be learned from the program.

SCORING AND ANALYSIS

The answer keys for the two forms are provided below:

Item No.	Form A	Form B
1	F	T
2	F	F
3	F	F
4	T	F
5	F	T
6	T	F
7	F	F
8	T	T
9	F	F
10	T	F
11	T	F
12	T	T
13	F	T
14	T	T
15	T	T

The measures should be scored by counting the number of correct answers for each participant. Items marked "Don't Know" or left blank should be scored as incorrect. Count the number of correct answers for each participant. Next, total the correct answers for the group and divide by the number of participants in the group. The mean number of correct answers and the standard deviation can be used to summarize participant performance on the measure. Means and standard deviations from before and after the program can be compared to determine changes in participants' knowledge.

EXERCISE FACTS

Form A

This test has 15 sentences about exercise. Put a check to show whether you think each sentence is TRUE or FALSE. If you don't know whether a sentence is true or false, put a check under DON'T KNOW.

True	False	Don't Know	
( )	( )	( )	1. People can tell if someone is fit just by looking at that person.
( )	( )	( )	2. Strength exercises are a good way to improve the heart and lungs.
( )	( )	( )	3. Regular exercise makes the heart pump less blood with each heartbeat.
( )	( )	( )	4. Regular exercise helps a person keep from getting fat.
( )	( )	( )	5. Exercising even once will make a person's heart stronger.
( )	( )	( )	6. Regular stretching can help a person feel less muscle stiffness.
( )	( )	( )	7. People who are thin don't need to exercise.
( )	( )	( )	8. Regular stretching can make muscles more flexible.
( )	( )	( )	9. Muscles that are not exercised turn into fat.
( )	( )	( )	10. Regular exercise helps keep the blood vessels from getting blocked.
( )	( )	( )	11. Regular exercise can help keep the breathing system healthy.
( )	( )	( )	12. When people exercise, they lose fat and increase muscle strength.
( )	( )	( )	13. People who exercise are more likely to have heart attacks than people who don't.
( )	( )	( )	14. To help the heart and lungs, exercise must be done every day.
( )	( )	( )	15. Regular exercise can make bones stronger.

EXERCISE FACTS

Form B

This test has 15 sentences about exercise. Put a check to show whether you think each sentence is TRUE or FALSE. If you don't know whether a sentence is true or false, put a check under DON'T KNOW.

True	False	Don't Know		
( )	( )	( )	1.	A person must exercise regularly to be physically fit.
( )	( )	( )	2.	People should exercise as hard as they can if they want it to be good for them.
( )	( )	( )	3.	Regular exercise makes the heart beat faster even when a person is not exercising.
( )	( )	( )	4.	People should not exercise while dieting because exercise makes people want to eat more.
( )	( )	( )	5.	Exercise can help people relax.
( )	( )	( )	6.	People usually lose a lot of weight when they first start exercising.
( )	( )	( )	7.	All sports have the same effects on the body.
( )	( )	( )	8.	Exercise can help a person avoid back pains.
( )	( )	( )	9.	Exercising a part of the body is a good way to reduce body fat in that part.
( )	( )	( )	10.	It is not important for people to exercise until they are 35 years old.
( )	( )	( )	11.	A person's weight is the best measure of how much fat the person has.
( )	( )	( )	12.	Muscles get smaller if they are not used.
( )	( )	( )	13.	Exercise can help people have lower blood pressure.
( )	( )	( )	14.	A person's heart rate must increase during exercise for it to have good results.
( )	( )	( )	15.	Exercise can help people meet new friends.

PLANNING A SAFE EXERCISE PROGRAM

(FORMS A & B)

This knowledge measure examines what participants know about planning and implementing a safe and effective exercise program. The measure is appropriate for adults.

PURPOSE

Information about participants' knowledge of planning and implementing a safe, effective exercise program may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show limited knowledge of a safe, effective exercise program, thus indicating a need for instruction in that area.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' knowledge in planning a safe, effective exercise program.

PROCEDURES

Because the equidifficulty of the forms has not been established, it is better not to give all participants Form A as a pretest and Form B as a posttest. Instead, choose either of the following methods.

- Review Forms A and B and select one. Give all participants the selected form both before and after the program. Alternatively, select 20 items from the two forms and construct a measure most consistent with your program emphasis. Then administer the "new" form both before and after the program.
- Give Form A to half of the incoming participants and Form B to the remaining half. To distribute the forms randomly, order them "ABABAB" and hand them out. Following the program, give each participant the form not previously taken. For example, if a participant was given Form B before the program, then that participant should be given Form A following the program. This approach eliminates the possibility that examinees will be sensitized to the specific facts to be learned from the program.

SCORING AND ANALYSIS

The answer keys for the two forms are provided below:

Item No.	Form A	Form B
1	F	F
2	T	T
3	F	F
4	T	T
5	F	F
6	T	F
7	T	T
8	F	F
9	T	F
10	T	F
11	T	T
12	F	T
13	F	F
14	F	F
15	T	T
16	F	T
17	T	F
18	T	F
19	F	T
20	F	T

The measures should be scored by counting the number of correct answers for each participant. Items marked "Don't Know" or left blank should be scored as incorrect. Count the number of correct answers for each participant. Next, total the correct answers for the group and divide by the number of participants in the group. The mean number of correct answers and the standard deviation can be used to summarize participant performance on the measure. Means and standard deviations from before and after the program can be compared to determine changes in participant knowledge.

PLANNING A SAFE EXERCISE PROGRAM

Form A

This test has 20 sentences about exercise programs. Put a check to show whether you think each sentence is TRUE or FALSE. If you don't know whether a sentence is true or false, put a check under DON'T KNOW.

True	**False**	**Don't Know**	
( )	( )	( )	1. A person's heart rate during exercise should not be greater than the person's heart rate at rest.
( )	( )	( )	2. One way to avoid heat disorders when exercising on hot days is to wear light, loose-fitting clothing.
( )	( )	( )	3. Individuals should start an exercise session with very vigorous activity.
( )	( )	( )	4. Lifting heavy weights may be dangerous for individuals with high blood pressure.
( )	( )	( )	5. All individuals should see a doctor before starting an exercise program.
( )	( )	( )	6. Individuals should seek medical attention if they feel any unusual pain while exercising.
( )	( )	( )	7. In warm weather, goose bumps on the upper arms and chest are an early warning sign of heat disorders.
( )	( )	( )	8. Bowling regularly is likely to improve heart/lung fitness.
( )	( )	( )	9. Tendinitis is a common tendon injury caused by exercise.
( )	( )	( )	10. Exercising in very cold weather can be dangerous.
( )	( )	( )	11. Yoga can improve muscular flexibility.
( )	( )	( )	12. A slow heartbeat is a symptom of altitude sickness.
( )	( )	( )	13. Most individuals with lung problems should avoid regular exercise.

True	False	Don't Know	
( )	( )	( )	14. Heart/lung fitness activities should maintain the heart rate at its target level for no more than 30 minutes.
( )	( )	( )	15. Cool-down exercises reduce the risk of muscle cramps.
( )	( )	( )	16. Body fluids lost during exercise can only be replaced with special salt solutions.
( )	( )	( )	17. The loss of body fluids is the major cause of heat disorders.
( )	( )	( )	18. Heat exhaustion is the first sign that the body is losing its ability to regulate temperature.
( )	( )	( )	19. Cool-down exercises should last at least 20 to 30 minutes.
( )	( )	( )	20. A person should lift heavy objects with the back and not with the legs.

PLANNING A SAFE EXERCISE PROGRAM

Form B

This test has 20 sentences about exercise programs. Put a check to show whether you think each sentence is TRUE or FALSE. If you don't know whether a sentence is true or false, put a check under DON'T KNOW.

True	False	Don't Know		
( )	( )	( )	1.	It is better for one's knees to run on hard surfaces than on soft surfaces such as dirt.
( )	( )	( )	2.	Bicycling can improve a person's heart/lung fitness.
( )	( )	( )	3.	People should wear a sweatshirt and sweatpants while exercising in hot weather.
( )	( )	( )	4.	If people feel pain in a joint while exercising, they should slow down or stop exercising.
( )	( )	( )	5.	Vegetarian diets do not provide enough nutrients for people who exercise regularly.
( )	( )	( )	6.	People should not drink large quantities of water while exercising.
( )	( )	( )	7.	People should not exercise as hard when smog levels are high.
( )	( )	( )	8.	Warm-up and stretching exercises frequently cause muscle strain.
( )	( )	( )	9.	The stretching period of an exercise session should consist of vigorous calisthenics.
( )	( )	( )	10.	Exercising in cold weather is generally a greater health risk than exercising in hot weather.
( )	( )	( )	11.	People who exercise in extreme cold should wear many layers of light clothing rather than one heavy garment.
( )	( )	( )	12.	It is safe for people with arthritis to exercise regularly.
( )	( )	( )	13.	A person's endurance is the same at all altitudes.

True	False	Don't Know	
( )	( )	( )	14. When people feel symptoms of heat stress, they should continue to exercise but with less intensity.
( )	( )	( )	15. Heat cramps are usually the result of water loss.
( )	( )	( )	16. People who experience extreme breathlessness after light exercise should have a medical check-up.
( )	( )	( )	17. A slow pulse is a sign of heat exhaustion.
( )	( )	( )	18. Everyone has the same target heart rate.
( )	( )	( )	19. An activity program where everyone exercises at the same intensity can be dangerous.
( )	( )	( )	20. An exercise program must allow time for muscles to recover after exercising.

PLANNING TO EXERCISE SAFELY

(FORMS A & B)

This knowledge measure examines what participants know about planning and implementing a safe, effective exercise program. The measure is appropriate for adolescents and preadolescents.

PURPOSE

Information about participants' knowledge of planning and implementing a safe, effective exercise program may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show limited knowledge of how to plan and implement a safe, effective exercise program, thus indicating a need for instruction in that area.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' knowledge of a safe, effective exercise program.

PROCEDURES

Because the equidifficulty of the forms has not been established, it is better not to give all participants Form A as a pretest and Form B as a posttest. Instead, choose either of the following methods.

- Review Forms A and B and select one. Give all participants the selected form both before and after the program. Alternatively, select 15 items from the two forms and construct a measure most consistent with your program emphasis. Then administer the "new" form both before and after the program.
- Give Form A to half of the incoming participants and Form B to the remaining half. To distribute the forms randomly, order them "ABABAB" and hand them out. Following the program, give each participant the form not previously taken. For example, if a participant was given Form B before the program, then that participant should be given Form A following the program. This approach eliminates the possibility that examinees will be sensitized to the specific facts to be learned from the program.

SCORING AND ANALYSIS

The answer keys for the two forms are provided below:

Item No.	Form A	Form B
1	F	F
2	F	F
3	T	T
4	T	T
5	T	F
6	T	T
7	F	F
8	F	F
9	F	T
10	T	T
11	F	F
12	T	T
13	T	F
14	F	F
15	T	T

The measures should be scored by counting the number of correct answers for each participant. Items marked "Don't Know" or left blank should be scored as incorrect. Count the number of correct answers for each participant. Next, total the correct answers for the group and divide by the number of participants in the group. The mean number of correct answers and the standard deviation can be used to summarize participant performance on the measure. Means and standard deviations from before and after the program can be compared to determine changes in participant knowledge.

PLANNING TO EXERCISE SAFELY

Form A

This test has 15 sentences about planning for exercise. Put a check to show whether you think each sentence is TRUE or FALSE. If you don't know whether a sentence is true or false, put a check under DON'T KNOW.

True	False	Don't Know		
( )	( )	( )	1.	Taking special medicines is a good way to improve one's exercise ability.
( )	( )	( )	2.	Everyone should see a doctor before starting an exercise program.
( )	( )	( )	3.	Exercising in very cold weather can be dangerous.
( )	( )	( )	4.	A complete physical fitness program should have heart/lung, flexibility, muscular strength, and muscular endurance exercises.
( )	( )	( )	5.	Young people should not lift heavy weights until their bones stop growing.
( )	( )	( )	6.	Warm-up exercises help the body get ready to do heart/lung exercises.
( )	( )	( )	7.	Cool-down exercises often cause muscle soreness.
( )	( )	( )	8.	Playing softball is a good way to improve heart/lung fitness.
( )	( )	( )	9.	It is best to wear the same shoes for all sports or exercises.
( )	( )	( )	10.	Swimming can improve muscular flexibility.
( )	( )	( )	11.	People should not drink water while exercising.
( )	( )	( )	12.	The heart/lung fitness period of an exercise program should last at least 20 minutes.
( )	( )	( )	13.	People should see a doctor if they feel any unusual pain while exercising.
( )	( )	( )	14.	People should exercise until they are completely tired.
( )	( )	( )	15.	Bicycling, swimming, and rope skipping can improve heart/lung fitness.

PLANNING TO EXERCISE SAFELY

Form B

This test has 15 sentences about planning for exercise. Put a check to show whether you think each sentence is TRUE or FALSE. If you don't know whether a sentence is true or false, put a check under DON'T KNOW.

True	False	Don't Know	
( )	( )	( )	1. It is better to run on hard surfaces than on soft surfaces like dirt.
( )	( )	( )	2. Exercising when one is sick will help a person get well faster.
( )	( )	( )	3. The difficulty of heart/lung fitness exercises can be checked by measuring the heart rate.
( )	( )	( )	4. A person can develop muscular endurance by lifting light weights many times in a row.
( )	( )	( )	5. A slow heart rate is a sign of heat problems.
( )	( )	( )	6. Heart/lung exercises must be done at least 3 times a week to improve heart/lung fitness.
( )	( )	( )	7. People who have not been exercising should start by exercising as hard as they can.
( )	( )	( )	8. Only adults over 30 years old need to do warm-up exercises before hard exercise.
( )	( )	( )	9. People should do flexibility exercises every day.
( )	( )	( )	10. The body cools itself by sweating.
( )	( )	( )	11. Exercising at one's maximum heart rate is safe for most people.
( )	( )	( )	12. People should exercise less on smoggy days than they usually do.
( )	( )	( )	13. To get stronger, people must lift weights every day.
( )	( )	( )	14. Cool-down exercises should last about 20 minutes.
( )	( )	( )	15. Stretching before and after exercising will help keep muscles from getting sore.

SELECTING AN EXERCISE PROGRAM

(FORMS A & B)

This skill measure assesses participants' ability to select an appropriate physical fitness program. This measure is appropriate for adults.

PURPOSE

Information about participants' ability to select an exercise program to meet personal fitness goals may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show a lack in participants' ability select appropriate exercise programs, thus indicating a need for participant training in that area.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' skills in selecting appropriate exercise programs.

PROCEDURES

Because the equidifficulty of the forms has not been established, it is better not to give all participants Form A as a pretest and Form B as a posttest. Instead, choose either of the following methods.

- Review Forms A and B and select one. Give all participants the selected form both before and after the program. Alternatively, select 10 items from the two forms and construct a measure most consistent with your program emphasis. Then administer the "new" form both before and after the program.
- Give Form A to half of the incoming participants and Form B to the remaining half. To distribute the forms randomly, order them "ABABAB" and hand them out. Following the program, give each participant the form not previously taken. For example, if a participant was given Form B before the program, then that participant should be given Form A following the program. This approach eliminates the possibility that examinees will remember how they answered each item from the pretest.

SCORING AND ANALYSIS

The answer keys for the two forms are provided below:

Item No.	Form A	Form B
1	B	D
2	D	A
3	A	B
4	A	B
5	C	C
6	D	A
7	B	D
8	C	C
9	A	B
10	C	C

The measures should be scored by counting the number of correct answers for each participant. Blank items should be scored as incorrect. Count the number of correct answers for each participant. Next, total the correct answers for the group and divide by the number of participants in the group. The mean number of correct answers and the standard deviation can be used to summarize participant performance on the measure. Means and standard deviations from before and after the program can be compared to determine changes in participant skill.

SELECTING AN EXERCISE PROGRAM

Form A

This test presents descriptions of individuals who want to select an exercise program that will meet their personal fitness goals. These individuals have no major medical problems. Put a check in the box under the plan that is most appropriate to achieve the fitness goal given.

1. **Name:** Ray Grand

Age: 29

Fitness goal: cardiorespiratory endurance and muscular strength

Put a check in the box under the plan that is most appropriate for Ray:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Volleyball	Jogging	Rowing	Swimming
Number of sessions per week:	3	3	2	4
Length of each session:	30 min.	20 min.	20 min.	30 min.
Target heart rate (for 1 minute):	108-128	136-150	174-192	180-196
Other activities:	Stretching	Weight training	Weight training	–
(*Check one box*)	☐	☐	☐	☐

2. **Name:** Donna White

Age: 58

Fitness goal: cardiorespiratory endurance and flexibility

Put a check in the box under the plan that is most appropriate for Donna:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Bicycling	Tennis	Brisk walking	Swimming
Number of sessions per week:	2	3	2	4
Length of each session:	20 min.	30 min.	15 min.	25 min.
Target heart rate (for 1 minute):	156-177	139-145	98-108	120-126
Other activities:	Golf	–	Yoga	Calisthenics

(*Check one box*) ☐ ☐ ☐ ☐

3. **Name:** Larry Jones

Age: 40

Fitness goal: cardiorespiratory endurance and weight control

Put a check in the box under the plan that is most appropriate for Larry:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Brisk walking	Calisthenics	Jogging	Golf (no cart)
Number of sessions per week:	5	7	2	1
Length of each session:	45 min.	30 min.	15 min.	90 min.
Target heart rate (for 1 minute):	136-144	98-108	162-170	90-102
Other activities:	–	–	–	–

(*Check one box*) ☐ ☐ ☐ ☐

4. **Name:** Gary Craft

Age: 17

Fitness goal: cardiorespiratory endurance

Put a check in the box under the plan that is most appropriate for Gary:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Racquetball	Basketball	Weight training	Golf (no cart)
Number of sessions per week:	3	2	3	1
Length of each session:	30 min.	45 min.	30 min.	90 min.
Target heart rate (for 1 minute):	145-160	122-145	162-170	90-102
Other activities:	–	–	–	–
(Check one box)	☐	☐	☐	☐

5. **Name:** Gloria Mays

Age: 49

Fitness goal: cardiorespiratory endurance and weight control

Put a check in the box under the plan that is most appropriate for Gloria:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Jogging	Swimming	Rope skipping	Brisk walking
Number of sessions per week:	1	6	5	2
Length of each session:	45 min.	10 min.	20 min.	20 min.
Target heart rate (for 1 minute):	174-180	126-144	122-132	138-145
Other activities:	–	–	–	–
(Check one box)	☐	☐	☐	☐

6. **Name:** Howard Ramos

Age: 42

Fitness goal: cardiorespiratory endurance and muscular strength

Put a check in the box under the plan that is most appropriate for Howard:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Badminton	Brisk walking	Bowling	Rowing
Number of sessions per week:	2	2	3	4
Length of each session:	30 min.	25 min.	35 min.	20 min.
Target heart rate (for 1 minute):	126-138	108-120	96-108	125-140
Other activities:	–	–	Weight training	–
(Check one box)	☐	☐	☐	☐

7. **Name:** Sylvia Rapp

Age: 48

Fitness goal: cardiorespiratory endurance and muscle strength

Put a check in the box under the plan that is most appropriate for Sylvia:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Brisk walking	Jogging	Calisthenics	Swimming
Number of sessions per week:	2	3	4	3
Length of each session:	20 min.	30 min.	15 min.	25 min.
Target heart rate (for 1 minute):	105-112	123-138	90-102	96-108
Other activities:	Bowling	Weight training	Golf	Weight training
(Check one box)	☐	☐	☐	☐

8. **Name:** Andrew Walters

Age: 62

Fitness goal: cardiorespiratory endurance

Put a check in the box under the plan that is most appropriate for Andrew:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Squash	Calisthenics	Bicycling	Bowling
Number of sessions per week:	1	2	3	5
Length of each session:	50 min.	35 min.	25 min.	45 min.
Target heart rate (for 1 minute):	125-1323	96-108	112-130	90-102
Other activities:	–	–	–	–
(Check one box)	☐	☐	☐	☐

9. **Name:** Marion Ward

Age: 39

Fitness goal: cardiorespiratory endurance and muscle endurance

Put a check in the box under the plan that is most appropriate for Marion:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Dancing (aerobic)	Skiing (cross-country)	Hiking	Handball
Number of sessions per week:	4	2	2	1
Length of each session:	30 min.	35 min.	40 min.	60 min.
Target heart rate (for 1 minute):	138-153	140-156	120-126	162-168
Other activities:	–	Yoga	–	Weight training
(Check one box)	☐	☐	☐	☐

10. **Name:** Marcia Michaels

Age: 19

Fitness goal: cardiorespiratory endurance and flexibility

Put a check in the box under the plan that is most appropriate for Marcia:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Skating	Weight training	Squash	Karate
Number of sessions per week:	2	3	4	3
Length of each session:	20 min.	25 min.	40 min.	25 min.
Target heart rate (for 1 minute):	120-126	102-114	143-156	96-108
Other activities:	Stretching	Yoga	Calisthenics	–
(*Check one box*)	☐	☐	☐	☐

SELECTING AN EXERCISE PROGRAM

Form B

This test presents descriptions of individuals who want to select an exercise program that will meet their personal fitness goals. These individuals have no major medical problems. Put a check in the box under the plan that is most appropriate to achieve the fitness goal given.

1. **Name:** Barry Osborne

Age: 58

Fitness goal: cardiorespiratory endurance and weight control

Put a check in the box under the plan that is most appropriate for Ray:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Soccer	Rope skipping	Softball	Brisk walking
Number of sessions per week:	1	6	3	5
Length of each session:	60 min.	5 min.	45 min.	30 min.
Target heart rate (for 1 minute):	150-159	125-140	90-102	115-130
Other activities:	–	Weight training	Calisthenics	–
(Check one box)	☐	☐	☐	☐

2. **Name:** Terry Jenson

Age: 28

Fitness goal: cardiorespiratory endurance and muscular strength

Put a check in the box under the plan that is most appropriate for Terry:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Rowing	Swimming	Basketball	Weight training
Number of sessions per week:	4	1	2	3
Length of each session:	20 min.	50 min.	30 min.	20 min.
Target heart rate (for 1 minute):	136-154	173-179	120-136	96-108
Other activities:	–	–	Calisthenics	–
(*Check one box*)	☐	☐	☐	☐

3. **Name:** Rachael Alberts

Age: 34

Fitness goal: cardiorespiratory endurance and muscular endurance

Put a check in the box under the plan that is most appropriate for Rachael:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Judo	Swimming	Dancing (social)	Tennis
Number of sessions per week:	4	3	2	1
Length of each session:	30 min.	30 min.	45 min.	75 min.
Target heart rate (for 1 minute):	96-108	136-148	132-150	150-162
Other activities:	–	–	–	–
(*Check one box*)	☐	☐	☐	☐

4. **Name:** Charles Jager

Age: 41

Fitness goal: cardiorespiratory endurance and muscular endurance

Put a check in the box under the plan that is most appropriate for Charles:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Weight training	Basketball	Racquetball	Badminton
Number of sessions per week:	4	3	2	2
Length of each session:	20 min.	30 min.	40 min.	20 min.
Target heart rate (for 1 minute):	90-102	125-140	162-178	114-144
Other activities:	–	–	–	–

(Check one box) ☐ ☐ ☐ ☐

5. **Name:** Luis Carlos

Age: 17

Fitness goal: cardiorespiratory endurance

Put a check in the box under the plan that is most appropriate for Luis:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Badminton	Calisthenics	Handball	Rowing
Number of sessions per week:	1	3	4	3
Length of each session:	40 min.	20 min.	20 min.	10 min.
Target heart rate (for 1 minute):	138-156	90-102	143-160	186-198
Other activities:	–	–	–	–

(Check one box) ☐ ☐ ☐ ☐

6. **Name:** Julie Mitchell

Age: 36

Fitness goal: cardiorespiratory endurance and weight control

Put a check in the box under the plan that is most appropriate for Julie:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Squash	Skating	Karate	Basketball
Number of sessions per week:	5	3	5	3
Length of each session:	30 min.	15 min.	20 min.	10 min.
Target heart rate (for 1 minute):	130-140	96-108	90-102	132-150
Other activities:	–	–	–	–
(*Check one box*)	☐	☐	☐	☐

7. **Name:** Gwen Harrington

Age: 20

Fitness goal: cardiorespiratory endurance and flexibility

Put a check in the box under the plan that is most appropriate for Gwen:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Yoga	Swimming	Rope skipping	Running
Number of sessions per week:	3	3	6	4
Length of each session:	30 min.	12 min.	5 min.	20 min.
Target heart rate (for 1 minute):	90-102	176-187	162-170	143-153
Other activities:	–	–	–	–
(*Check one box*)	☐	☐	☐	☐

8. **Name:** Lydia McClain

Age: 40

Fitness goal: cardiorespiratory endurance and weight control

Put a check in the box under the plan that is most appropriate for Lydia:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Brisk walking	Handball	Bicycling	Swimming
Number of sessions per week:	4	2	4	3
Length of each session:	20 min.	40 min.	25 min.	10 min.
Target heart rate (for 1 minute):	114-120	140-156	136-150	158-170
Other activities:	–	–	–	–
(Check one box)	☐	☐	☐	☐

9. **Name:** Sherri Collins

Age: 44

Fitness goal: cardiorespiratory endurance

Put a check in the box under the plan that is most appropriate for Sherri:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Yoga	Dancing (aerobic)	Rope skipping	Golf (no cart)
Number of sessions per week:	3	4	2	3
Length of each session:	20 min.	30 min.	15 min.	45 min.
Target heart rate (for 1 minute):	[illegible]	[illegible]	[illegible]	[illegible]
Other activities:	–	–	–	–
(Check one box)	☐	☐	☐	☐

10. **Name:** Arthur Wingerski

Age: 33

Fitness goal: cardiorespiratory endurance and flexibility

Put a check in the box under the plan that is most appropriate for Arthur:

	Plan A	Plan B	Plan C	Plan D
Cardiorespiratory activity:	Jogging	Brisk walking	Rowing	Basketball
Number of sessions per week:	2	2	3	3
Length of each session:	20 min.	20 min.	20 min.	25 min.
Target heart rate (for 1 minute):	140-150	90-102	140-150	120-132
Other activities:	Wrestling	Karate	Yoga	Weight training
(Check one box)	☐	☐	☐	☐

PREVENTING AND CARING FOR INJURIES

(FORMS A & B)

This skill measure assesses participants' ability to prevent and care for exercise-related injuries appropriately. This measure is appropriate for adults.

PURPOSE

Information on participants' ability to prevent and care for exercise-related injuries may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show that participants have limited skills in appropriately preventing and caring for exercise-related injuries, thus indicating a need for participant training in that area.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' skills in preventing and caring for exercise-related injuries.

PROCEDURES

Because the equidifficulty of the forms has not been established, it is better not to give all participants Form A as a pretest and Form B as a posttest. Instead, choose either of the following methods.

- Review Forms A and B and select one. Give all participants the selected form both before and after the program. Alternatively, select 15 items from the two forms and construct a measure most consistent with your program emphasis. Then administer the "new" form both before and after the program.
- Give Form A to half of the incoming participants and Form B to the remaining half. To distribute the forms randomly, order them "ABABAB" and hand them out. Following the program, give each participant the form not previously taken. For example, if a participant was given Form B before the program, then that participant should be given Form A following the program. This approach eliminates the possibility that examinees will remember how they answered each item from the pretest.

SCORING AND ANALYSIS

The answer keys for the two forms are provided below:

Item No.	Form A	Form B
1	B	D
2	A	D
3	C	C
4	B	B
5	D	B
6	A	A
7	D	B
8	C	B
9	B	C
10	C	D
11	A	A
12	C	C
13	B	B
14	A	A
15	C	D

The measures should be scored by counting the number of correct answers for each participant. Blank items should be scored as incorrect. Count the number of correct answers for each participant. Next, total the correct answers for the group and divide by the number of participants in the group. The mean number of correct answers and the standard deviation can be used to summarize participant performance on the measure. Means and standard deviations from before and after the program can be compared to determine changes in participant skill.

PREVENTING AND CARING FOR INJURIES

Form A

This test presents descriptions of individuals who want to prevent or care for exercise-related injuries. Circle the letter of the appropriate action for the individual to take. If there is no choice presented that is appropriate, circle Choice D, NONE OF THE ABOVE.

1. Jane had no difficulty sleeping before she started a regular exercise program. However, now she has trouble sleeping. She would like to take care of her sleep problem. An appropriate action for Jane to take would be to:
   A. Increase the intensity and length of her exercise program.
   B. Reduce the intensity and length of her exercise program.
   C. Avoid eating for at least two hours before exercising.
   D. None of the above.

2. Brian has begun a running program and wants to reduce the risk of getting tendinitis in his ankles. An appropriate action for Brian to take would be to:
   A. Perform warm-up exercises before running.
   B. Perform cool-down exercises after running.
   C. Run in an area that is free of obstacles and traffic.
   D. None of the above.

3. Harvey wants to avoid exercise-related heart and lung problems as he begins a bicycling program. An appropriate action for Harvey to take would be to:
   A. Use only a stationary bicycle at first.
   B. Avoid bicycling right before any meal.
   C. Gradually build up his fitness level.
   D. None of the above.

4. Anita has started a dance program. She feels a sore spot developing on her foot. She discovers she has a water blister. She wants to take care of her blister. An appropriate action for Anita to take would be to:
   A. Soak her feet in warm water.
   B. Clean her foot, then cover the blister with a bandage.
   C. Take aspirin unless advised not to by a physician.
   D. None of the above.

5. Bill has just seriously sprained his ankle while playing basketball. He wants to act safely. An appropriate action for Bill to take would be to:
   A. Apply hot compresses to his ankle.
   B. Sit down but continue to slowly move his ankle.
   C. Wrap his ankle and continue playing with caution.
   D. None of the above.

6. Corrine is beginning a rowing program. She wants to prevent heat problems. An appropriate action for Corrine to take would be to:
   A. Wear light, loose clothing, and drink a lot of water.
   B. Take salt tablets and perform cool-down exercises after rowing.
   C. Wear waterproof clothing and perform warm-up exercises before rowing.
   D. None of the above.

7. Hilary is starting an ice skating class. She wants to prevent losing a lot of body heat. An appropriate action for Hilary to take would be to:
   A. Perform warm-up exercises before each class.
   B. Increase the length of the cool-down period.
   C. Drink a hot beverage after class.
   D. None of the above.

8. Tina's toes became frostbitten while she was hiking. She wants to take care of the frostbite. An appropriate action for Tina to take would be to:
   - A. Apply ointment and bandages to her toes.
   - B. Put her feet in cold water.
   - C. Put her feet in lukewarm water.
   - D. None of the above.

9. Terry plays handball regularly. He is suffering from tendinitis in his shoulder. He wants to relieve the symptoms. An appropriate action for Terry to take would be to:
   - A. Tape his shoulder right before playing.
   - B. Apply ice to his shoulder immediately after playing.
   - C. Apply warm towels to his shoulder immediately after playing
   - D. None of the above.

10. Gary wants to avoid fracturing any bones while playing football. An appropriate action for Gary to take would be to:
    - A. Exercise with weights before each game to develop his muscular strength.
    - B. Increase the amount of time he spends cooling down after a game.
    - C. Use the appropriate protective equipment and know the rules of the game.
    - D. None of the above.

11. Sandra plays racquetball regularly. She has a fever from a virus and wants to reduce the risk of getting a viral infection of the heart muscle. An appropriate action for Sandra to take would be to:
    - A. Stop exercising until several days after she has fully recovered
    - B. Play racquetball for only half the usual length of time while she is recovering.
    - C. Play racquetball to try to "sweat" out the fever.
    - D. None of the above.

12. Jose, a man in his early 50s, is bicycling in the afternoon with his friends. He begins to feel dizzy, weak, and out of breath, and his pulse is very rapid. An appropriate action for Jose to take would be to:
    A. Rest until his pulse is normal and then ride home.
    B. Drink plenty of liquids before riding home.
    C. Seek medical care.
    D. None of the above.

13. Andrea is in her 20s and wants to start swimming regularly. She wants to act safely. An appropriate action for Andrea to take would be to:
    A. Have a medical check-up before starting to swim.
    B. Gradually increase how long and how hard she swims each time.
    C. Avoid eating for two hours after she finishes a swim.
    D. None of the above.

14. Larry wants to avoid heat problems as he begins a tennis program. An appropriate action for Larry to take would be to:
    A. Drink plenty of water while playing.
    B. Play only in the early afternoon.
    C. Play in a sweatshirt and sweatpants.
    D. None of the above.

15. Barney used to jog regularly in high school. Now, a few years later, he intends to start a regular jogging program again. He wants to avoid injury. An appropriate action for Barney to take would be to:
    A. Start his program at the same pace as before.
    B. Avoid jogging at the end of the day.
    C. Begin the program at a comfortable level.
    D. None of the above.

PREVENTING AND CARING FOR INJURIES

Form B

This test presents descriptions of individuals who want to prevent or care for exercise-related injuries. Circle the letter of the appropriate action for the individual to take. If there is no choice presented that is appropriate, circle Choice D, NONE OF THE ABOVE.

1. Phillip is starting to play handball regularly. He wants to avoid heart/lung problems. An appropriate action for Phillip to take would be to:

   A. Play with a partner who is an advanced player.

   B. Play on indoor rather than outdoor courts.

   C. Take lessons to improve his ability.

   D. None of the above.

2. Janet is beginning a regular jogging program. She wants to prevent muscle strains. An appropriate action for Janet to take would be to:

   A. Perform cool-down exercises after jogging.

   B. Jog only on a concrete running surface.

   C. Jog every day.

   D. None of the above.

3. Suzanne began a bicycling program two weeks ago. Recently, she has felt nauseous after exercising. She wants to act safely. An appropriate action for Suzanne to take would be to:

   A. Perform warm-up exercises for a longer time before bicycling.

   B. Increase the intensity but reduce the amount of time she bicycles.

   C. Reduce the intensity of the bicycling and increase the cool-down period.

   D. None of the above.

4. Art is going cross-country skiing for the afternoon. He wants to prevent getting dangerously cold. An appropriate action for Art to take would be to:
   A. Wear a thick layer of loosely woven clothing, protective goggles, and knitted gloves.
   B. Wear several thick layers of tightly woven clothing, a hat, and mittens.
   C. Perform warm-up and cool-down exercises.
   D. None of the above.

5. Shari has been jogging regularly for several months. Lately she has noticed bursts of rapid heartbeats while jogging. She wants to act safely. An appropriate action for Shari to take would be to:
   A. Stop jogging for several days, then resume jogging at a slower pace.
   B. Seek medical care.
   C. Reduce the intensity and the length of her jogging.
   D. None of the above.

6. Louise is playing tennis on a warm summer day. She begins to suffer from a heat problem. She wants to act safely. An appropriate action for Louise to take would be to:
   A. Stop playing tennis, rest in a cooler location, and drink water.
   B. Rest while putting an ice pack on her head.
   C. Continue playing in a cooler location.
   D. None of the above.

7. Frank wants to prevent dislocating his shoulder while playing football. An appropriate action for Frank to take would be to:
   A. Perform warm-up exercises for 15 to 20 minutes before the game starts.
   B. Use appropriate protective clothing and equipment and proper playing techniques.
   C. Perform cool-down exercises for 15 to 20 minutes after the game is over.
   D. None of the above.

8. Martha is 46 years old. She is playing badminton with her friend. She starts to feel painful pressure in her chest. An appropriate action for Martha to take would be to:
   A. Continue to play, but with less intensity.
   B. Seek medical care.
   C. Stop playing badminton and do some cool-down exercises.
   D. None of the above.

9. Jim is suffering from tendinitis in his elbow from playing tennis. He wants to relieve the symptoms. An appropriate action for Jim to take would be to:
   A. Apply heat to his elbow after playing.
   B. Massage his elbow before playing.
   C. Apply ice to his elbow after playing.
   D. None of the above.

10. Carlos becomes extremely breathless while playing basketball. His breathlessness lasts for 15 to 20 minutes after he stops playing. He wants to act safely. An appropriate action for Carlos to take would be to:
    A. Continue to play hard but shorten his playing time.
    B. Lengthen his warm-up period.
    C. Lengthen his cool-down period.
    D. None of the above.

11. Gina jumps rope regularly. She wants to reduce the risk of getting tendinitis in her ankles. An appropriate action for Gina to take would be to:
    A. Keep the amount of time and intensity of jumping rope at a comfortable level.
    B. Perform cool-down exercises after she jumps rope.
    C. Tape her ankles with elastic bandages.
    D. None of the above.

12. Rose sprained her ankle during a gymnastics class. She wants to act safely. An appropriate action for Rose to take would be to:
    - A. Massage her ankle gently.
    - B. Sit with her ankle raised and apply hot towels to it.
    - C. Sit with her ankle raised and apply ice to it.
    - D. None of the above.

13. Michael wants to avoid heart/lung problems as he begins a swimming program. An appropriate action for Michael to take would be to:
    - A. Swim only in a heated pool.
    - B. Slowly build up his fitness level.
    - C. Take swimming lessons to improve his form.
    - D. None of the above.

14. Juan has a fever from a viral illness. He hikes a great deal. Juan wants to reduce the risk of getting a viral infection of the heart muscle. An appropriate action for Juan to take would be to:
    - A. Stop hiking until he is completely recovered from his illness.
    - B. Take only short hiking trips until he is completely recovered.
    - C. Stop hiking but lift weights to keep up his strength.
    - D. None of the above.

15. Renee is going to take a walk through snowy mountains. She wants to prevent frostbite. An appropriate action for Renee to take would be to:
    - A. Walk at a slow, easy pace.
    - B. Wear a heavy layer of loosely woven clothes.
    - C. Walk in the middle of the day.
    - D. None of the above.

EXERCISING SAFELY

(FORMS A & B)

This skill measure assesses participants' ability to prevent and care for exercise-related injuries appropriately. This measure is appropriate for adolescents and preadolescents.

PURPOSE

Information on participants' ability to prevent and care for exercise-related injuries may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show participants have limited skills in appropriately preventing and caring for exercise-related injuries, thus indicating a need for participant training in that area.
- When the measure is administered prior to and following a program, it is possible to evaluate changes in participants' skills in preventing and caring for exercise-related injuries.

PROCEDURES

Because the equidifficulty of the forms has not been established, it is better not to give all participants Form A as a pretest and Form B as a posttest. Instead, choose either of the following methods.

- Review Forms A and B and select one. Give all participants the selected form both before and after the program. Alternatively, select 10 items from the two forms and construct a measure most consistent with your program emphasis. Then administer the "new" form both before and after the program.
- Give Form A to half of the incoming participants and Form B to the remaining half. To distribute the forms randomly, order them "ABABAB" and hand them out. Following the program, give each participant the form not previously taken. For example, if a participant was given Form B before the program, then that participant should be given Form A following the program. This approach eliminates the possibility that examinees will remember how they answered each item from the pretest.

SCORING AND ANALYSIS

The answer keys for the two forms are provided below:

Item No.	Form A	Form B
1	A	C
2	B	A
3	B	B
4	C	B
5	A	A
6	B	C
7	B	B
8	C	C
9	C	A
10	A	B

The measures should be scored by counting the number of correct answers for each participant. Blank items should be scored as incorrect. Count the number of correct answers for each participant. Next, total the correct answers for the group and divide by the number of participants in the group. The mean number of correct answers and the standard deviation can be used to summarize participant performance on the measure. Means and standard deviations from before and after the program can be compared to determine changes in participant skill.

EXERCISING SAFELY

Form A

These questions are about young people who want to prevent or take care of an exercise injury. For each question, circle the letter of the best action for the person to take.

1. Janet rides a skateboard. She wants to prevent scrapes. The best action for Janet to take is to:
   - A. Wear protective clothing.
   - B. Do warm-up and cool-down exercises.
   - C. Use her skateboard on the sidewalk only.

2. Jimmy is going to be in a bicycle race with his friends next month. He wants to do as well as he can without getting hurt. The best action for Jimmy to take is to:
   - A. Take steroids to build up his leg muscles.
   - B. Ride his bike several times a week.
   - C. Lift heavy weights to increase his strength.

3. Leon's toes became very cold and painful while he was playing in the snow. He wants to take care of his toes. The best action for Leon to take is to:
   - A. Drink warm liquids.
   - B. Soak his toes in cool water.
   - C. Put his toes in the hottest water that he can stand.

4. Sally wants to take care of a blister on her hand. The best action for Sally to take is to:
   - A. Soak her hand in hot water for several minutes.
   - B. Put ice on the blister for several minutes.
   - C. Clean her hand, then cover the blister with a bandage.

5. Jeff pitches for his baseball team. Sometimes his wrist swells up. He wants to prevent swelling of his wrist. The best action for Jeff to take is to:
   A. Do warm-up exercises for his wrist before playing.
   B. Put ice on his wrist before playing.
   C. Rub his wrist during the game.

6. Ray is bicycling with his friends. He begins to feel a bad pain in his chest. He wants to act safely. The best action for Ray to take is to:
   A. Ride home as fast as he can.
   B. Stop riding and have a friend telephone for help.
   C. Continue riding at a slower speed.

7. Kevin is going to play on the school football team. He wants to avoid breaking any bones while playing football. The best action for Kevin to take is to:
   A. Play less hard during practice than he does during the game.
   B. Know the rules of the game and wear protective equipment.
   C. Do warm-up exercises for one hour before each game.

8. Juan swims regularly. He has started feeling very tired after swimming. He wants to act safely. The best action for Juan to take is to:
   A. Swim more often to get in better shape.
   B. Swim in colder water.
   C. Swim at a slower speed.

9. Clara is taking a dance-exercise class. She wants to keep from twisting her ankle. The best action for Clara to take is to:
   A. Do cool-down exercises after each class.
   B. Always do the exercises with a partner.
   C. Learn how to do the exercises correctly.

10. Martha strained her shoulder while she was playing softball. She wants to care for her shoulder before going to the doctor. The best action for Martha to take is to:
   A. Put ice on her shoulder.
   B. Gently rub her shoulder.
   C. Wrap her shoulder in bandages.

EXERCISING SAFELY

Form B

These questions are about young people who want to prevent or take care of an exercise injury. For each question circle the letter of the best action for the person to take.

1. Cindy is going to run in a Jog-a-thon to help raise money for her school. She wants to run as many laps as possible in 20 minutes without getting hurt. The best action for Cindy to take is to:
   - A. Eat some candy for quick energy one-half hour before running.
   - B. Wear new running shoes for the Jog-a-thon.
   - C. Jog several times a week for at least a month before the Jog-a-thon.

2. Larry started swimming every day. Lately, he has had an upset stomach while he swims. He wants to act safely. The best action for Larry to take is to:
   - A. Avoid eating for two hours before swimming.
   - B. Drink some milk right before he swims to coat his stomach.
   - C. Avoid eating for two hours after swimming.

3. Anita is going hiking on a very warm day. She wants to prevent heat problems. The best action for Anita to take is to:
   - A. Walk very quickly so that she will feel a breeze.
   - B. Wear loose, light clothing and drink plenty of water.
   - C. Do cool-down exercises as soon as the hike is over.

4. Mickey is learning how to play baseball. He wants to keep from getting scraped when he slides into the bases. The best action for Mickey to take is to:
   - A. Wear shoes made especially for baseball.
   - B. Wear protective clothing.
   - C. Do warm-up exercises before playing.

5. Karen plays soccer. She wants to keep from twisting her ankles. The best action for Karen to take is to:
   - A. Play in a clear area and wear correctly fitted shoes.
   - B. Perform cool-down exercises at the end of the game.
   - C. Wear ankle guards while she is playing.

6. Eric is starting to play a great deal of basketball. He wants to prevent muscle strains. The best action for Eric to take is to:
   - A. Spend less time warming-up before games.
   - B. Avoid playing on outdoor basketball courts.
   - C. Do stretching exercises before each game.

7. Gary has a blister on his heel from running. He wants to take care of the blister. The best action for Gary to take is to:
   - A. Put ice on the blister for several minutes.
   - B. Clean his heel, then put a bandage on the blister.
   - C. Soak the blister in warm water for several minutes.

8. Cecilia is going to play tennis on Saturday. She wants to keep from getting sunburned. The best action for Cecilia to take is to:
   - A. Play only in the early afternoon.
   - B. Wear sunglasses while she plays.
   - C. Wear a hat and put on some sun screen.

9. Ray scraped his knees while roller skating. He wants to care for his knees. The best action for Ray to take is to:
   - A. Wash his knees.
   - B. Let his knees get better without doing anything.
   - C. Put hot towels on his knees.

10. Jerome sprained his wrist while playing handball. He wants to take care of his wrist. The best action for Jerome to take is to:
    - A. Wrap his wrist in warm bandages.
    - B. Hold his wrist still and put ice on it.
    - C. Gently wiggle his wrist and fingers.

IDEAS ABOUT DECISIONS

This affective measure assesses participants' belief in the value of careful decision making. This measure is appropriate for adolescents and preadolescents.

PURPOSE

Information about decision making may be useful for the following reasons:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may indicate a need for strengthening participants' appreciation for careful decision making.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' beliefs regarding careful decision making.

PROCEDURES

This instrument can be administered both at the beginning and at the end of the program. However, handbook users should be alert to concerns regarding the potential reactivity of affective measures. A measure is considered *reactive* if the experience of completing the measure prior to the program causes participants to react differently to the program. Handbook users should, therefore, carefully review each affective measure that they wish to use to determine its potential for making participants unduly sensitive to aspects of the program. If a measure is determined to be reactive, then program personnel should *not* administer that measure to *all* participants as a pretest and posttest. Instead, the measure could be administered to half of the program participants prior to program participation to determine participants' pre-program status. The measure could then be administered to the other half of the participants after program participation to assess participants' post-program status.

SCORING AND ANALYSIS

Point values are assigned to responses as follows:

Item No.	Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
1	5	4	3	2	1
2	1	2	3	4	5
3	5	4	3	2	1
4	1	2	3	4	5
5	1	2	3	4	5
6	1	2	3	4	5
7	1	2	3	4	5
8	5	4	3	2	1
9	5	4	3	2	1
10	5	4	3	2	1

This inventory can be scored by adding the point values of the responses from all participants and dividing this total by the number of responses. Blank items should *not* be counted in the number of responses. The maximum attainable score of 5 points indicates a strong belief in the utility of making decisions carefully. A minimum score of 1 indicates a weak belief in the utility of making decisions carefully.

IDEAS ABOUT DECISIONS

The sentences below are about making decisions. For each sentence, place a check to show how much you agree or disagree with the sentence.

	Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
1. It is worth the time it takes to make decisions carefully.	( )	( )	( )	( )	( )
2. People should go with their first ideas when making decisions.	( )	( )	( )	( )	( )
3. People are happier with their decisions when they take the time to make them carefully.	( )	( )	( )	( )	( )
4. Spending a lot of time to make careful decisions is too difficult.	( )	( )	( )	( )	( )
5. Making careful decisions takes too much time.	( )	( )	( )	( )	( )
6. When making decisions, people should do what they feel, not what they think.	( )	( )	( )	( )	( )
7. People make equally good decisions no matter how they arrive at them.	( )	( )	( )	( )	( )
8. People who make quick decisions are usually disappointed with them later.	( )	( )	( )	( )	( )
9. People should take time to make decisions carefully.	( )	( )	( )	( )	( )
10. It is easy to make decisions carefully.	( )	( )	( )	( )	( )

EFFECTS OF EXERCISE

This affective measure assesses participants' belief regarding the possible effects of exercise on a person's body image, self-concept, ability to manage stress, and health. This measure is appropriate for adults.

PURPOSE

Information about participants' belief in the possible effects of exercise may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show that participants have little belief in the positive effects of exercise, thus indicating a need for instruction in that area.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' beliefs about the positive effects of exercise.

PROCEDURES

This instrument can be administered both at the beginning and at the end of the program. However, handbook users should be alert to concerns regarding the potential reactivity of affective measures. A measure is considered *reactive* if the experience of completing the measure prior to the program causes participants to react differently to the program. Handbook users should, therefore, carefully review each affective measure that they wish to use to determine its potential for making participants unduly sensitive to aspects of the program. If a measure is determined to be reactive, then program personnel should *not* administer that measure to *all* participants as a pretest and posttest. Instead, the measure could be administered to half of the program participants prior to program participation to determine participants' pre-program status. The measure could then be administered to the other half of the participants after program participation to assess participants' post-program status.

SCORING AND ANALYSIS

Point values are assigned to responses as follows:

Strongly Agree = 5
Agree = 4
Not Sure = 3
Disagree = 2
Strongly Disagree = 1

This inventory can be scored by adding the point values of the responses from all participants and dividing this total by the number of responses. Blank items should not be counted in the number of responses. The maximum attainable score of 5 points indicates a strong belief in the positive effects of exercise. A minimum score of 1 indicates weak belief in the positive effects of exercise.

EFFECTS OF EXERCISE

This survey describes some possible effects of regular exercise. For each statement place a check to show how much you agree or disagree with the statement.

	Strongly Agree	**Agree**	**Not Sure**	**Disagree**	**Strongly Disagree**
1. Exercise can help me control my weight.	( )	( )	( )	( )	( )
2. I am more productive when I exercise.	( )	( )	( )	( )	( )
3. Exercise can help me stay healthy.	( )	( )	( )	( )	( )
4. Exercise can help me meet people.	( )	( )	( )	( )	( )
5. I feel better about myself when I exercise.	( )	( )	( )	( )	( )
6. Exercise can make me more attractive.	( )	( )	( )	( )	( )
7. I get sick less often when I exercise than when I don't.	( )	( )	( )	( )	( )
8. Exercise can help me reduce stress.	( )	( )	( )	( )	( )
9. Exercise can help me live longer.	( )	( )	( )	( )	( )
10. I have more energy when I exercise than when I don't.	( )	( )	( )	( )	( )

BELIEFS ABOUT EXERCISE

This affective measure assesses participants' beliefs in the possible physiological, psychological, and sociological effects of exercise. This measure is appropriate for adolescents and preadolescents.

PURPOSE

Information about participants' beliefs in the possible positive effects of exercise may be useful for two reasons:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may indicate little belief in the positive effects of exercise and thus indicate a need for participant training in that area.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' beliefs about the positive effects of exercise.

PROCEDURES

This instrument can be administered both at the beginning and at the end of the program. However, handbook users should be alert to concerns regarding the potential reactivity of affective measures. A measure is considered *reactive* if the experience of completing the measure prior to the program causes participants to react differently to the program. Handbook users should, therefore, carefully review each affective measure that they wish to use to determine its potential for making participants unduly sensitive to aspects of the program. If a measure is determined to be reactive, then program personnel should *not* administer that measure to *all* participants as a pretest and posttest. Instead, the measure could be administered to half of the program participants prior to program participation to determine participants' pre-program status. The measure could then be administered to the other half of the participants after program participation to assess participants' post-program status.

SCORING AND ANALYSIS

Point values are assigned to responses as follows:

Strongly Agree = 5
Agree = 4
Not Sure = 3
Disagree = 2
Strongly Disagree = 1

This inventory can be scored by adding the point values of the responses from all participants and dividing this total by the number of responses. Blank items should not be counted in the number of responses. The maximum attainable score of 5 points indicates a strong belief in the positive effects of exercise on a variety of factors. A minimum score of 1 indicates weak belief in the positive effects of exercise across a variety of factors.

BELIEFS ABOUT EXERCISE

The sentences below are about exercise. For each sentence place a check to show how much you agree or disagree with the sentence.

	Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
1. Exercise can help keep me from worrying.	( )	( )	( )	( )	( )
2. Exercise can help me stay healthy.	( )	( )	( )	( )	( )
3. Exercise can help me make friends.	( )	( )	( )	( )	( )
4. Exercise can help me live longer.	( )	( )	( )	( )	( )
5. I have more energy when I exercise than when I don't.	( )	( )	( )	( )	( )
6. I feel better about myself when I exercise.	( )	( )	( )	( )	( )
7. I get sick less often when I exercise than when I don't.	( )	( )	( )	( )	( )
8. I look better when I exercise.	( )	( )	( )	( )	( )
9. Exercise can help me control my weight.	( )	( )	( )	( )	( )
10. I study and work better when I exercise.	( )	( )	( )	( )	( )

EXERCISING REGULARLY

This affective measure assesses participants' expectation to exercise regularly in situations when people might typically avoid exercising. This measure is appropriate for adults.

PURPOSE

Information about participants' expectation to exercise regularly may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show that participants have a low expectation to exercise regularly in situations not conducive to exercise, thus indicating a need for instruction in the positive effects of exercise.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' expectation to exercise regularly.

PROCEDURES

This instrument can be administered both at the beginning and at the end of the program. However, handbook users should be alert to concerns regarding the potential reactivity of affective measures. A measure is considered *reactive* if the experience of completing the measure prior to the program causes participants to react differently to the program. Handbook users should, therefore, carefully review each affective measure that they wish to use to determine its potential for making participants unduly sensitive to aspects of the program. If a measure is determined to be reactive, then program personnel should *not* administer that measure to *all* participants as a pretest and posttest. Instead, the measure could be administered to half of the program participants prior to program participation to determine participants' pre-program status. The measure could then be administered to the other half of the participants after program participation to assess participants' post-program status.

SCORING AND ANALYSIS

Point values are assigned to responses as follows:

Definitely Yes = 5
Probably Yes = 4
Maybe = 3
Probably No = 2
Definitely No = 1

This inventory can be scored by adding the point values of the responses from all participants and dividing this total by the number of responses. Blank items should not be counted in the number of responses. The maximum attainable score of 5 points indicates a high expectation to exercise regularly in a variety of situations. A minimum score of 1 indicates a low expectation to exercise regularly in a variety of situations not conducive to exercising.

EXERCISING REGULARLY

This survey is about exercising regularly in different situations. Regular exercise requires 20 minutes or more of planned activity at least three times per week.

Place a check to show how likely you would be to *exercise regularly* in the situation described in each question.

Would you exercise regularly if ...	**Definitely Yes**	**Probably Yes**	**Maybe**	**Probably No**	**Definitely No**
1. you were on a vacation?	( )	( )	( )	( )	( )
2. the places to exercise were inconvenient?	( )	( )	( )	( )	( )
3. you were very busy?	( )	( )	( )	( )	( )
4. your regular exercise partners decided to quit for awhile?	( )	( )	( )	( )	( )
5. you moved to a new neighborhood?	( )	( )	( )	( )	( )
6. you could not spend money to exercise?	( )	( )	( )	( )	( )
7. friends or relatives were staying with you for several weeks?	( )	( )	( )	( )	( )
8. you were tired?	( )	( )	( )	( )	( )
9. the weather was bad?	( )	( )	( )	( )	( )
10. you were overworked?	( )	( )	( )	( )	( )
11. you had just recovered from an injury you got while exercising?	( )	( )	( )	( )	( )

Would you exercise regularly if . . .	Definitely Yes	Probably Yes	Maybe	Probably No	Definitely No
12. you had been feeling depressed for quite a while?	( )	( )	( )	( )	( )
13. you had not been exercising and were out of condition?	( )	( )	( )	( )	( )
14. you had not made as much progress in exercising as you expected?	( )	( )	( )	( )	( )
15. you were happy with the amount you weighed?	( )	( )	( )	( )	( )
16. your spouse/friend did not exercise?	( )	( )	( )	( )	( )

INTENTION TO EXERCISE

This affective measure assesses participants' plans to begin or maintain regular exercise. This measure is appropriate for adults.

If this measure seems useful, you might also want to consider the **Weekly Activities Index** which is a behavior measure that examines participants' actual amount of regular exercise.

PURPOSE

Information about participants' intention to exercise may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show participants have little intention of exercising, thus indicating a need for instruction in the advantages of regular exercise.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' intention to exercise regularly.

PROCEDURES

This instrument can be administered both at the beginning and at the end of the program. However, handbook users should be alert to concerns regarding the potential reactivity of affective measures. A measure is considered *reactive* if the experience of completing the measure prior to the program causes participants to react differently to the program. Handbook users should, therefore, carefully review each affective measure that they wish to use to determine its potential for making participants unduly sensitive to aspects of the program. If a measure is determined to be reactive, then program personnel should *not* administer that measure to *all* participants as a pretest and posttest. Instead, the measure could be administered to half of the program participants prior to program participation to determine participants' pre-program status. The measure could then be administered to the other half of the participants after program participation to assess participants' post-program status.

SCORING AND ANALYSIS

Point values are assigned to responses as follows:

Definitely Yes = 5
Probably Yes = 4
Maybe = 3
Probably No = 2
Definitely No = 1

This inventory can be scored by adding the point values of the responses from all participants and dividing this total by the number of responses. Blank items should not be counted in the number of responses. The maximum attainable score of 5 points indicates a strong intention to continue exercising or to begin exercising regularly over a variety of time frames. A minimum score of 1 indicates a weak intention to continue exercising or to begin exercising over a variety of time frames.

INTENTION TO EXERCISE

This survey asks about your plans to begin or maintain regular exercise. Regular exercise requires 20 minutes or more of planned activity three times a week.

First, check one box to indicate whether you exercise regularly. Then, use the following scale to answer the questions under the box you check.

A	**B**	**C**	**D**	**E**
Definitely Yes	**Probably Yes**	**Maybe**	**Probably No**	**Definitely No**

CHECK ONE BOX:

I exercise at least 20 minutes 3 times per week.

☐ **Yes** ↓

1. Will you exercise regularly throughout the next week? (Circle one)

   A B C D E

2. Will you exercise regularly throughout the next three months? (Circle one)

   A B C D E

3. Will you exercise regularly throughout the next six months? (Circle one)

   A B C D E

☐ **No** ↓

1. Do you intend to begin exercising within the next week? (Circle one)

   A B C D E

2. Do you intend to begin exercising regularly within the next month? (Circle one)

   A B C D E

ATTITUDE TOWARD WORK

This affective measure assesses participants' attitude toward their present work situation. This measure is appropriate for adults.

PURPOSE

Many experts believe that regular exercise will improve one's attitude toward work. When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' attitudes toward work.

PROCEDURES

This instrument can be administered both at the beginning and at the end of the program. However, handbook users should be alert to concerns regarding the potential reactivity of affective measures. A measure is considered *reactive* if the experience of completing the measure prior to the program causes participants to react differently to the program. Handbook users should, therefore, carefully review each affective measure that they wish to use to determine its potential for making participants unduly sensitive to aspects of the program. If a measure is determined to be reactive, then program personnel should *not* administer that measure to *all* participants as a pretest and posttest. Instead, the measure could be administered to half of the program participants prior to program participation to determine participants' pre-program status. The measure could then be administered to the other half of the participants after program participation to assess participants' post-program status.

SCORING AND ANALYSIS

Point values are assigned to responses as follows:

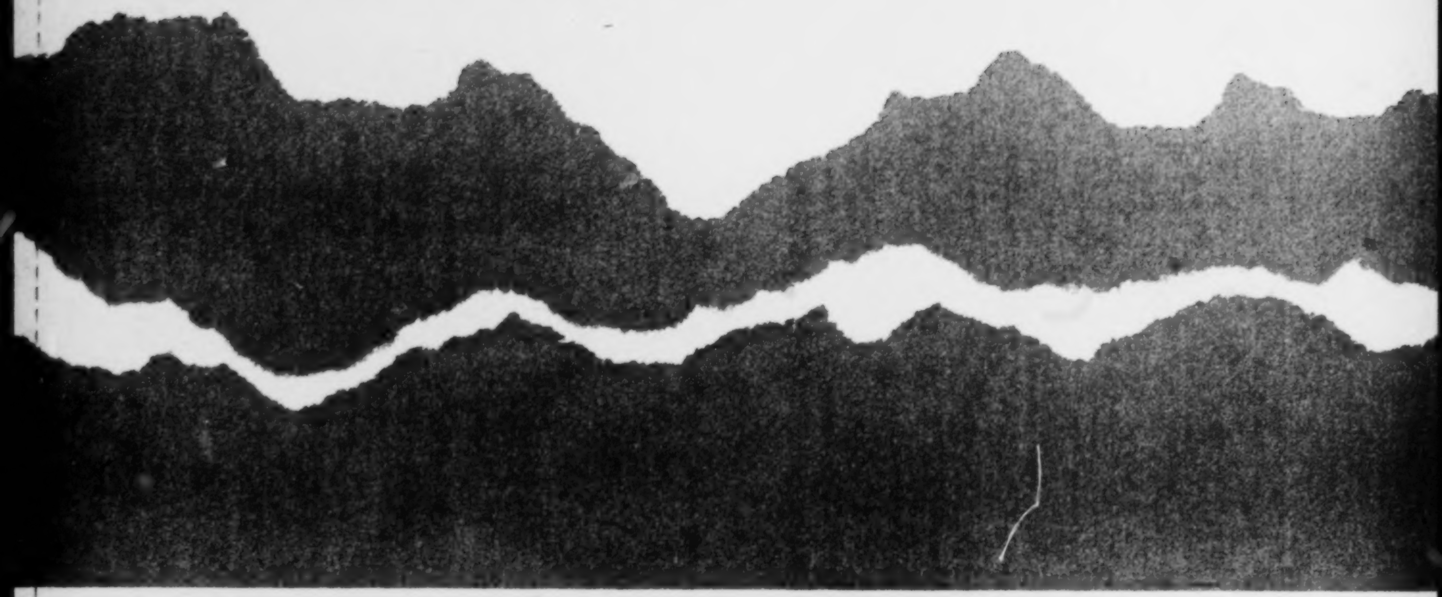
Item No.	Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
1	1	2	3	4	5
2	5	4	3	3	2
3	1	2	3	4	4
4	1	2	3	4	5
5	5	4	3	2	1
6	5	4	3	2	1
7	5	4	3	2	1
8	5	4	3	2	1
9	5	4	3	2	1
10	5	4	3	2	1

This inventory can be scored by adding the point values of the responses from all participants and dividing this total by the number of responses. Blank items should not be counted in the number of responses. The maximum attainable score of 5 points indicates a positive attitude toward work. A minimum score of 1 indicates a negative attitude toward work.

ATTITUDE TOWARD WORK

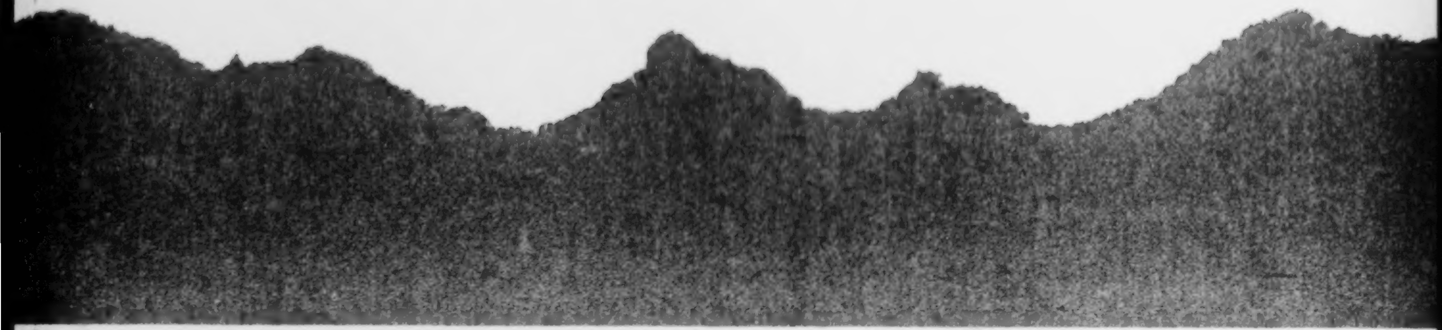
This survey asks you about your present work situation. Place a check to show how much you agree or disagree with each statement.

	Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
1. I have trouble concentrating at work.	( )	( )	( )	( )	( )
2. I can handle the pressures of work.	( )	( )	( )	( )	( )
3. I feel tired and worn out at work.	( )	( )	( )	( )	( )
4. I have to force myself to go to work.	( )	( )	( )	( )	( )
5. I am very productive at work.	( )	( )	( )	( )	( )
6. I think that my work is of good quality.	( )	( )	( )	( )	( )
7. I get along well with my co-workers.	( )	( )	( )	( )	( )
8. I am satisfied with my opportunities for promotion.	( )	( )	( )	( )	( )
9. I am satisfied with my opportunities to develop new skills.	( )	( )	( )	( )	( )
10. On the whole, I am satisfied with my job.	( )	( )	( )	( )	( )



CHAPTER FOUR

Physiological Measures



Considerations for Physiological Testing

The fitness testing measures included in this handbook can be used to evaluate program participants' physical fitness status at the beginning of a program as well as monitor changes in their physical fitness following a program. Unlike a medical evaluation, however, fitness testing cannot provide individual participants with medical clearance to engage in physical fitness programs, nor can it provide information to diagnose physical abnormalities. Instead, these measures should be used to assist in program planning and evaluation, which, in turn, can lead to more effective and comprehensive exercise programs.

The physiological measures included in this chapter were selected based on the following criteria: safety, range of application, simplicity of administration and interpretation, minimal equipment expenditure, and proven validity and reliability. Each fitness measure is introduced with a description of (a) the test's evaluation capacity, (b) the impact of the fitness component being tested on the participants' overall health, (c) the equipment needed to perform the test, (d) administrative suggestions regarding the implementation of the test, (e) step-by-step administration procedures, (f) information for scoring and analyzing the results, (g) existing technical information pertaining to the instrument, and (h) the norms and standards currently used to determine participants' fitness classification.

Use of measures

Handbook users may use any of the measures in this chapter without seeking further permission. The measures are presented in the sequence in which they should generally be administered. For example, the PAR-Q questionnaire should always be administered first to identify individuals requiring physician's approval to participate in a fitness testing program. Evaluator discretion should be used when determining which measures would be most appropriate to use with adolescents.

Testing considerations

Program personnel should seriously consider obtaining informed consent from program participants prior to any fitness testing. Informed consent ensures that participants are knowledgeable about the testing procedures and the associated physical demands before they agree to participate. A description of the procedures for obtaining informed consent can be found in Appendix B.

To obtain the most accurate testing results, it is suggested that the following factors be considered prior to testing:

- The time of day affects many of the physiological factors that determine cardiorespiratory functioning. Therefore, the time of day should be noted and remain constant in test-retest evaluations.
- Ambient temperature, relative humidity, and air movement have an effect on how the available cardiac output is divided between the active muscles and the cutaneous blood vessels for cooling. Because this distribution affects participants' cardiovascular output, program

personnel should attempt to provide the following environmental considerations:

Temperature of 64° - 72°F
Relative humidity of 60 percent or less
Still air

- Food intake affects heart rate, blood sugar levels, and respiration. The heart rate may rise for an hour or more after the ingestion of food, while a complete fast may result in low blood sugar during testing. Therefore, participants should only eat a light meal at least one hour prior to test administration. Participants should also avoid drinking alcohol for at least six hours prior to testing and avoid smoking or drinking tea or coffee for at least two hours prior to testing.
- Participants should avoid strenuous activity on the day prior to testing and on the day of testing. A one-hour rest period prior to testing is recommended.

Safety Considerations

Prior to fitness testing, program personnel should consider the following safety precautions:

- Test administrators should have a written plan for handling emergencies.
- Elderly participants and participants with known or suspected heart problems or uncontrolled hypertension (high blood pressure) should not be tested without the approval of their physicians.
- Test administrators should be trained in CPR (Cardiopulmonary Resuscitation) and ECC (Emergency Cardiac Care).
- Participants should be encouraged to warm up before beginning fitness tests.

Overview of Physiological Measures

Category	Title	Target Group	Description	Page No.
Cardio-respiratory Function	**Physical Activity Readiness Questionnaire (PAR-Q)**	Adults	Identifies individuals for whom fitness testing might be inappropriate.	117
	Bruce Treadmill Test	Adults	Assesses cardiorespiratory endurance through maximum aerobic capacity.	119
	Bicycle Ergometer	Adults		123
	3-Minute Step Test	Adults		133
	Cooper's 1.5-Mile or 12-Minute Run/Walk	Adults Adolescents		137
	Distance Run	Preadolescents		140
Body Composition	**Jackson and Pollock's Skinfold Measures**	Adults	Assesses body composition based on subcutaneous adipose tissue.	143
	Sum of Skinfold Fat	Adolescents Preadolescents		151
	Waist Circumference	Adults	Assesses body composition based on deep adipose tissue.	156
	Buttocks (Hip) Circumference	Adults		158

Category	Title	Target Group	Description	Page No.
Muscular Strength and Endurance	**Push-ups (Canadian Fitness Test)**	Adults	Assesses upper body strength and endurance.	160
	Fitnessgram® Flexed-arm Hang	Adolescents Preadolescents		163
	Sit-ups	Adults	Assesses abdominal strength and endurance.	166
	Modified Sit-ups	Adolescents Preadolescents		169
Flexibility	**Sit and Reach Test**	Adults	Assesses flexibility of the low back and posterior thighs.	172
	AAHPERD Sit and Reach Test	Adolescents Preadolescents		176

PHYSICAL ACTIVITY READINESS QUESTIONNAIRE (PAR-Q)

This measure evaluates participants' current physical fitness as it relates to their readiness to participate in a cardiorespiratory fitness assessment. This measure is appropriate for adults.

PURPOSE

Sudden and unaccustomed vigorous exercise can be dangerous for some participants with certain medical conditions. The PAR-Q identifies the small number of adults for whom vigorous physical activity might be inappropriate or those who should have medical advice concerning the type of activity most suitable for them. This instrument should be administered prior to fitness assessments.

PROCEDURE

Hand out a questionnaire to each participant. Have participants carefully read and complete the questionnaire.

SCORING AND ANALYSIS

Participants who answer yes to any of the questions should not participate in a physical fitness assessment until a physician has thoroughly examined them for cardiovascular abnormalities or other problems.

TECHNICAL INFORMATION

According to a study of 1130 adult office workers conducted by Shepard, Cox, & Simper (1981), the eight participants who were medically advised not to take the fitness tests were also identified by the PAR-Q; thus, the sensitivity rating for this test was 100%. Because there was also a large number of false positives (18.5%), the specificity rating was 81.4%.

Overall, Shepard and colleagues concluded that "although the PAR-Q cannot predict the more subtle and rare exercise-induced ECG changes that may carry an increased risk of cardiac catastrophe, the PAR-Q can discern which individuals should avoid vigorous exercise as reliably as a brief medical examination." They also conclude that the PAR-Q seems preferable to the more expensive alternative of medical supervision of all fitness tests.

Physical Activity Readiness Questionnaire (PAR-Q)*

PARTICIPANT IDENTIFICATION

PAR Q & YOU

PAR-Q is designed to help you help yourself. Many health benefits are associated with regular exercise, and the completion of PAR-Q is a sensible first step to take if you are planning to increase the amount of physical activity in your life.

For most people physical activity should not pose any problem or hazard. PAR-Q has been designed to identify the small number of adults for whom physical activity might be inappropriate or those who should have medical advice concerning the type of activity most suitable for them.

Common sense is your best guide in answering these few questions. Please read them carefully and check (√) the ☐ YES or ☐ NO opposite the question if it applies to you.

YES	NO	
☐	☐	1. Has your doctor ever said you have heart trouble?
☐	☐	2. Do you frequently have pains in your heart and chest?
☐	☐	3. Do you often feel faint or have spells of severe dizziness?
☐	☐	4. Has a doctor ever said your blood pressure was too high?
☐	☐	5. Has your doctor ever told you that you have a bone or joint problem such as arthritis that has been aggravated by exercise, or might be made worse with exercise?
☐	☐	6. Is there a good physical reason not mentioned here why you should not follow an activity program even if you wanted to?
☐	☐	7. Are you over age 65 and not accustomed to vigorous exercise?

If You Answered

YES to one or more questions

If you have not recently done so, consult with your personal physician by telephone or in person BEFORE increasing your physical activity and/or taking a fitness test. Tell him what questions you answered YES on PAR-Q, or show him your copy.

programs'

After medical evaluation, seek advice from your physician as to your suitability for:

- unrestricted physical activity, probably on a gradually increasing basis.
- restricted or supervised activity to meet your specific needs, at least on an initial basis. Check in your community for special programs or services

NO to all questions

If you answered PAR-Q accurately, you have reasonable assurance of your present suitability for:

- A GRADUATED EXERCISE PROGRAM - A gradual increase in proper exercise promotes good fitness development while minimizing or eliminating discomfort.
- AN EXERCISE TEST - Simple tests of fitness (such as the Canadian Home Fitness Test) or more complex types may be undertaken if you so desire.

postpone

If you have a temporary minor illness, such as a common cold.

---

* Developed by the British Columbia Ministry of Health. Conceptualized and critiqued by the Multidisciplinary Advisory Board on Exercise (MABE). Translation, reproduction and use in its entirety is encouraged. Modifications by written permission only. Not to be used for commercial advertising in order to solicit business from the public.
Reference: PAR-Q Validation Report, British Columbia Ministry of Health, 1978.

* Produced by the British Columbia Ministry of Health and the Department of National Health & Welfare

BRUCE TREADMILL TEST

This measure assesses the maximum aerobic capacity on a treadmill. This measure is appropriate for adults.

PURPOSE

Because the treadmill test allows use of the familiar movements of running and walking, it is one of the most widely used tests of cardiorespiratory function. This test involves a slightly larger muscle mass than the bicycle ergometer test and results in a five to eight percent increase in maximum oxygen consumption, or VO_2 max, without the localized fatigue of the quadriceps for the amount of work done. It is also the only test that involuntarily controls the rate of energy expenditure (American Heart Association, 1972). Information about participants' maximum aerobic capacity may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show that participants have a low aerobic capacity, thus indicating the need for a cardiovascular training program.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' maximum aerobic capacity.

ADMINISTRATIVE SUGGESTIONS

- People with severe chest pain should **NOT** be tested.
- On the day prior to testing and on the day of testing, participants should refrain from any physical exertion.
- Participants should eat a *light* meal at least one hour before testing.
- Note the time of day that the test is administered so that it may be a constant in a test-retest situation.
- Attach at least seven electrodes to the chest for adequate monitoring.
- Anyone administering this test should have special training regarding this test and have a background in CPR.
- The test should begin gradually to allow participants to warm up, and it should wind down gradually to allow participants to cool down.
- Encourage participants to use the handrail at the beginning and end of the test for safety.
- The endpoint is determined individually when fatigue and/or other limiting symptoms and signs appear. Mandatory indications for stopping the test should be the appearance of an ataxic or

uncoordinated gait or of three or more consecutive ventricular premature beats.

- When the test is given by a non-physician, testing should be stopped if any of the following conditions appear: dizziness, angina, unusual or intolerable fatigue or intolerable pain. Some signs of intolerance are staggering or unsteadiness, mental confusion, facial expressions signifying disorders (strained or blank faces), pallor, rapid or distressful breathing, nausea or vomiting, and a definite fall in systolic blood pressure with increasing work load due to the heart's inability to contract as frequently or forcefully. If a physician is conducting the test, the physician may choose a different set of criteria.

PROCEDURE

For the high-fitness level participant. The test includes seven 3-minute stages. After an appropriate warm-up period, increase the speed and percent grade every three minutes as follows:

Stage	Speed (MPH)	Percent Grade
One	1.7	10
Two	2.5	12
Three	3.4	14
Four	4.2	16
Five	5.0	18
Six	5.5	20
Seven	6.0	22

For the low-fitness level participant. The test includes four 3-minute stages. After an appropriate warm-up period, increase the speed and percent grade every three minutes as follows:

Stage	Speed (MPH)	Percent Grade
One	1.2	0
Two	1.2	3
Three	1.2	6
Four	1.7	6

SCORING AND ANALYSIS

In this test, participants are assumed to have reached the limit of their oxygen uptake (VO_2 max). Within narrow limits, each work level requires a specific oxygen uptake per kilogram of body weight; therefore the VO_2 max for individuals may be closely estimated from the stage of the test at which they are forced to stop. For example, if a 35-year-old American male stopped at the end of Stage 3, he would have a VO_2 max of approximately 35 ml O_2 (oxygen) per kilogram, which is an average limit for a sedentary American male, aged 35. If this same man completed Stage 4, he would have a VO_2 max of 49 ml O_2 per kilogram, representing an above-average level of fitness for his age group.

Use Table 1 in the norms and standards section to determine the amount of oxygen needed for the test. Then compare participants' scores to the Cardiorespiratory Fitness Classification chart to determine their fitness category.

When the participant is allowed to begin training, an individualized program should be designed. Training should include cardiovascular exercises such as bench-stepping, running, rope-jumping, regular cycling, or swimming.

TECHNICAL INFORMATION

Lack of sufficient practice on the treadmill and holding on to the rail during testing may cause inaccurate measures for VO_2 max estimations. To increase sensitivity, several baseline electrocardiograms must be taken prior to beginning the treadmill phase of the test.

Studies have shown that the correlation between measured and estimated VO_2 max is approximately .94, which is considered acceptable in field studies with adult men and women (Matarazzo, 1983).

NORMS AND STANDARDS

TABLE 1

O_2 Requirement ml O_2/kg/min.	3-Minute Stages	
56.0		
52.5	mph	% gr
49.0		
45.5	4.2	16
42.0		
38.5		
35.0	3.4	14
31.5		
28.0		
24.5	2.5	12
21.0		
17.5	1.7	10
14.0		
10.5		
7.0		
3.5		

Cardiorespiratory Fitness Classification†

WOMEN

Maximal Oxygen Uptake (ml/kg/min)

Age (yrs)	Low	Fair	Average	Good	High
20-29	<24	24-30	31-37	38-48	49+
30-39	<20	20-27	28-33	34-44	45+
40-49	<17	17-23	24-30	31-41	42+
50-59	<15	15-20	21-27	28-37	38+
60-69	<13	13-17	18-23	24-34	35+

MEN

Maximal Oxygen Uptake (ml/kg/min)

Age (yrs)	Low	Fair	Average	Good	High
20-29	<25	25-33	34-42	43-52	53+
30-39	<23	23-30	31-38	39-48	49+
40-49	<20	20-26	27-35	36-44	45+
50-59	<18	18-24	25-33	34-42	43+
60-69	<16	16-22	23-30	31-40	41+

†Data from Prevention Medicine Center, Palo Alto, Calif, and from a survey of published sources.

Cited from American Heart Association's *Exercise testing and training of apparently healthy individuals: A handbook for physicians.*

BICYCLE ERGOMETER

This measure evaluates cardiovascular fitness by measuring the heart rate. This measure is appropriate for adults.

PURPOSE

The bicycle test predicts maximum working capacity or the response to submaximal work. It can also be used to predict maximum oxygen consumption, although it is not the bicycle ergometer's primary purpose. Information about participants' cardiorespiratory endurance may be useful in the following ways:

- Administration of this measure at the beginning of a program may provide needs assessment information. For example, results of this measure may show that participants have a low level of cardiorespiratory endurance, thus indicating a need for a cardiovascular training program.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' cardiorespiratory fitness.
- When used with first year fitness program participants, regular testing (once every 10-12 weeks) can demonstrate participants' response to training and possibly act as a motivator for continued participation. After the first year, testing should be done once a year.

EQUIPMENT

- An accurate, easily calibrated, constant torque bicycle ergometer with a range of 0-2100 kilogram-meters per minute is needed. Each major graduation should be at 300 kilogram-meters, with intermediate marks at 150 kilogram-meters.
- Charting graphs for each participant

ADMINISTRATIVE SUGGESTIONS

- Some individuals might be unfamiliar with bicycling. Therefore, a practice session prior to the actual testing is recommended.
- On the day of testing, participants should refrain from any physical exertion and should abstain from smoking or eating for two hours prior to the test.
- There is a linear relationship between heart rate (HR) and work; however, this linearity only exists at certain heart rates. At low heart rates, many external stimuli will affect the HR—talking, laughing, nervousness, etc. However, once the heart starts pumping harder and the muscles demand blood, external stimuli no longer affect the

---

Reprinted from *The Y's Way to Physical Fitness*, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

heart rate, and linearity occurs. This occurs at about 110 beats per minute (bpm). The relationship between HR and work increases in a linear fashion until it plateaus, signaling the maximum HR.

The purpose of this test is to establish the linearity between HR and work for the person being tested. To establish a line, two points are needed; therefore, two workloads are used. The only precaution is that these two points must be in the linear portion of the relationship. The workloads cannot be too high or too low. Linearity begins at approximately 110 bpm. The plateauing due to reaching maximum heart rate is a function of age; however, at a heart rate of 150 bpm almost everyone tested will be linear. Therefore, linearity is said to be between 110 and 150 bpm.

To eliminate the need to guess the workload needed to start the test, see the Guide To Setting Workloads on the Bicycle Ergometer. Use of the chart should eliminate the possibility of presenting too difficult a workload for a participant. The workload chart should be used conservatively as it is better to give a workload that is too low rather than one that is too high. The first workload is given to determine the heart rate response that is being elicited for a low workload. Usually this first workload will not be plotted because the HR will be under 110 bpm; however, should the HR be above 110 bpm, it should be used and then only one more workload will be necessary to plot the line. If the heart rate is not 110 bpm or greater, then two more workloads will be needed to plot points.

PROCEDURE

1. Check the calibration of the bicycle. See the Calibration Procedure for a simplified description of the calibration process. Because there is a slight difference in resistance between bicycles, be sure that any retesting of an individual is done on the same bicycle.
2. Briefly explain the test to the participants and have them fill out the top part of the individual record forms (Bicycle Ergometer Heart Rate Form and Prediction Graph). Copies of these forms that can be easily reproduced are provided at the end of this section.
3. Adjust the seat height. A participant's knee should be straight, with the ball of the foot on the pedal and the leg stretched. Record the seat position so that it can be used when retesting.
4. Set the metronome at 50 bpm and allow the participant to pedal freewheel (no load) for a minute to get the pace.
5. Set the first workload at 300 kgm/min (1.0 KP). Allow the participant to work at the first workload for three minutes. Count the heart rate between the second and third minutes. The difference in heart rates between the second and third minutes should not vary by more than five beats; if they do, extend the ride for an extra minute or until a stable value is obtained.

Reprinted from *The Y's Way to Physical Fitness*, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

Note: If the first workload elicits a HR of 110 or more, it should be plotted on the graph and only **ONE** more workload will be necessary. At both the second and third workloads, time should be taken at the end of the second and third minutes.

The Guide to Setting Workloads on the Bicycle Ergometer should be used only as a guide; hence, common sense should also be used. It is best to be conservative and use lower workloads for borderline scores.

The two plot points should be in the linear portion of the curve (approximately 110-150 bpm). It is better to have the two points toward the low end of this linearity.

An overweight and obviously unfit male might need to use the female scale for setting the workload; a trim, fit, young woman might be more appropriately paced according to the men's guidelines.

SCORING AND ANALYSIS

After the test is completed, the final heart rate in each of the workloads to be used (the two between 110 bpm and 150 bpm) should be plotted against the respective workload on the graph. Determine the participant's maximum heart rate by subtracting the participant's age from 220, then draw a horizontal line across the graph. Draw a straight line through the two points and extend the line to the participant's predicted maximal heart rate line. The point at which the diagonal line intersects the horizontal line (i.e., the predicted maximal heart rate line) represents the *maximal working capacity* for that participant. A perpendicular line should be drawn from this point to the baseline where the *maximal physical workload capacity* can be read in kgm/min.

The greatest source of error for the physical working capacity test is the possibility that the age-estimated maximum heart rate is not correct. Research has shown that maximum heart rates have a wide range at any age. Accuracy can be improved if the true maximum heart rate is known. This is usually not readily available. If the norm tables provided here are used, do **NOT** use the actual maximum heart rate, even if it is known. Instead use the formula 220 minus a participant's age for the maximum heart rate.

TECHNICAL INFORMATION

No matter how test results are used, the validity of the interpretation and its usefulness to a participant depends on the quality of the data. To ensure high-quality data, the test must be conducted as described. A bicycle ergometer must be well-maintained, which means regular calibration and proper maintenance. The environment must be well controlled; this implies freedom from both physical and emotional stress.

The mean difference between the physical working capacity determined by the bicycle ergometer and the actual physical working capacity was between .023 and .059 for men and .010 and .051 for women in liters of O_2 per minute. This results in a percent error of 6.7% for men and 9.4% for women (Astrand, 1954). The correlation between the bicycle ergometer and the step test is .95 (Kasch, 1966).

Reprinted from *The Y's Way to Physical Fitness*, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

NORMS AND STANDARDS

BICYCLE ERGOMETER

Physical Working Capacity

(Maximum kgm)

Females

Percentage Ranking	Age		
	18-35	36-45	46 +
95	1700	1600	1500
85	1500	1400	1300
75	1300	1200	1100
50	1100	1000	900
30	900	800	700
15	700	600	500

BICYCLE ERGOMETER

Physical Working Capacity

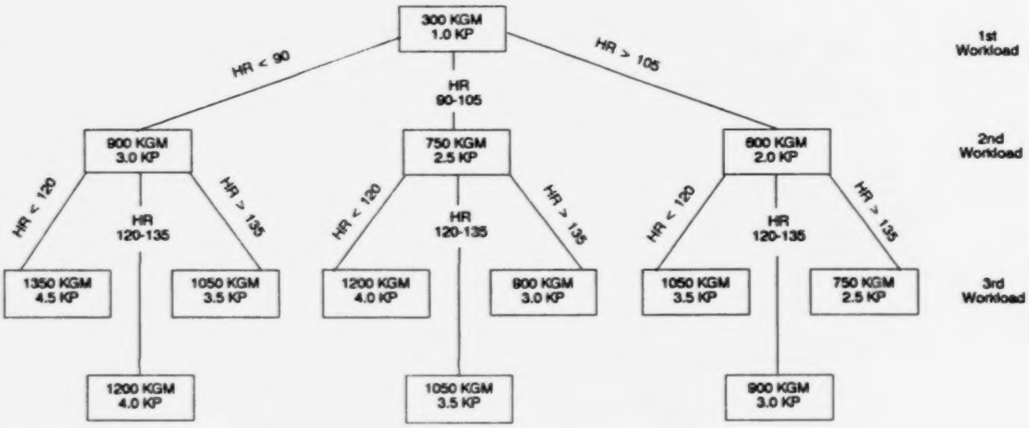
(Maximum kgm)

Males

Percentage Ranking	Age		
	18-35	36-45	46 +
95	2000	1800	1700
85	1800	1600	1500
75	1700	1500	1400
50	1500	1300	1200
30	1300	1100	1000
15	1200	1000	900

Reprinted from *The Y's Way to Physical Fitness*, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

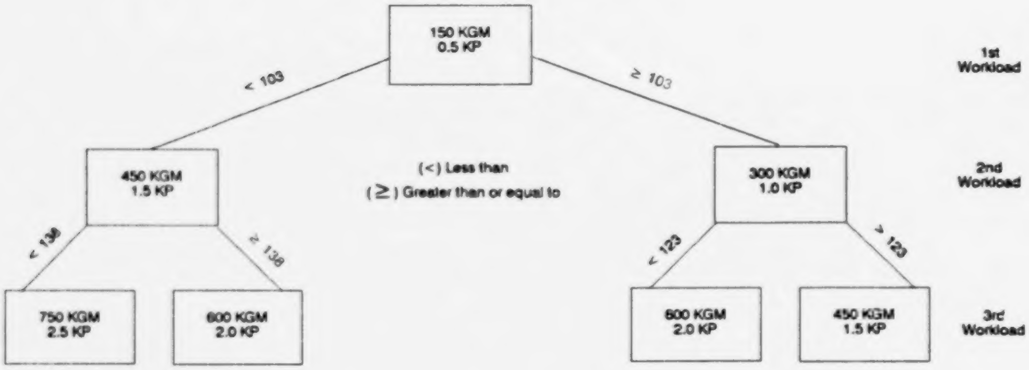
GUIDE TO SETTING WORKLOADS ON THE BICYCLE ERGOMETER
MALES


300 KGM
1.0 KP
1st Workload
HR < 90
HR 90-105
HR > 105
900 KGM
3.0 KP
750 KGM
2.5 KP
600 KGM
2.0 KP
2nd Workload
HR < 120
HR 120-135
HR > 135
HR < 120
HR 120-135
HR > 135
HR < 120
HR 120-135
HR > 135
1350 KGM
4.5 KP
1050 KGM
3.5 KP
1200 KGM
4.0 KP
900 KGM
3.0 KP
1050 KGM
3.5 KP
750 KGM
2.5 KP
3rd Workload
1200 KGM
4.0 KP
1050 KGM
3.5 KP
900 KGM
3.0 KP

DIRECTIONS

1. Set the 1st workload at 300 kgm/min (1.0 KP).
2. If HR in 3rd min is: Less than (<) 90, set 2nd load at 900 kgm (3 KP)
   Between 90 and 105, set 2nd load at 750 kgm (2.5 KP)
   Greater than (>) 105, set 2nd load at 600 km (2.0 KP)
3. Follow the same pattern for setting 3rd and final load.
4. NOTE: If the 1st workload elicits a HR of 110 or more, it is used on the graph, and only ONE more workload will be necessary.

FEMALES

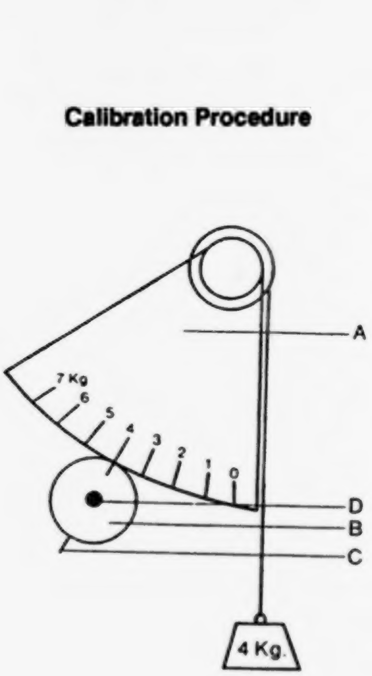

150 KGM
0.5 KP
1st Workload
< 103
≥ 103
450 KGM
1.5 KP
(<) Less than
(≥) Greater than or equal to
300 KGM
1.0 KP
2nd Workload
< 138
≥ 138
< 123
> 123
750 KGM
2.5 KP
600 KGM
2.0 KP
600 KGM
2.0 KP
450 KGM
1.5 KP
3rd Workload

DIRECTIONS

1. Set the 1st workload at 150 kgm/min (.5 KP).
2. If steady-state heart rate is < 103, set 2nd load at 450 kgm (1.5 KP)
   If steady-state heart rate is ≥ 103, set 2nd load at 300 kgm (1.0 KP)
3. Follow the same pattern for setting 3rd and final load.
4. NOTE: If the 1st workload elicits a HR of 110 or more, it is used on the graph, and only ONE more workload will be necessary.

Reprinted from *The Y's Way to Physical Fitness*, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

Calibration Procedure



The calibration of the bike is done precisely at the factory and unless the adjusting screw (C) has been tampered with, seldom is there a need for recalibration. However, incorrect calibration can be checked as follows:

Set the mark on the pendulum weight (B) at "0." Attach a weight known to be accurate as shown above. A 1 kg weight should correspond to a reading of 1 kg on the scale (A); a 2 kg weight should correspond to a reading of 2 kg on the scale (A); and so on. The example above shows 4 kg corresponding to 4 kg on the scale.

If the numbers do not match, make a correction by changing the adjusting screw (C). This screw moves the center of gravity of the pendulum [this screw is locked with the screw (D)].

Reprinted from *The Y's Way to Physical Fitness*, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

Bicycle Ergometer Heart Rate Form
BICYCLE ERGOMETER
Physical Work Capacity Test

Age 40 Weight 176 lbs. 80 kg.

Seat Height 8 Predicted Max Heart Rate 180 B/M

WORKLOADS			HEART RATE	
1st Workload 300 kgm				2nd min
			105	3rd min
				4th min (if needed)
2nd Workload	600	kgm	116	2nd min
			120	3rd min
				4th min
3rd Workload	900	kgm	145	2nd min
			145	3rd min
				4th min (if needed)

Reprinted from *The Y's Way to Physical Fitness*, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

Prediction Graph

Y's WAY TO PHYSICAL FITNESS — TEST BATTERY

MAXIMUM PHYSICAL WORKING CAPACITY PREDICTION

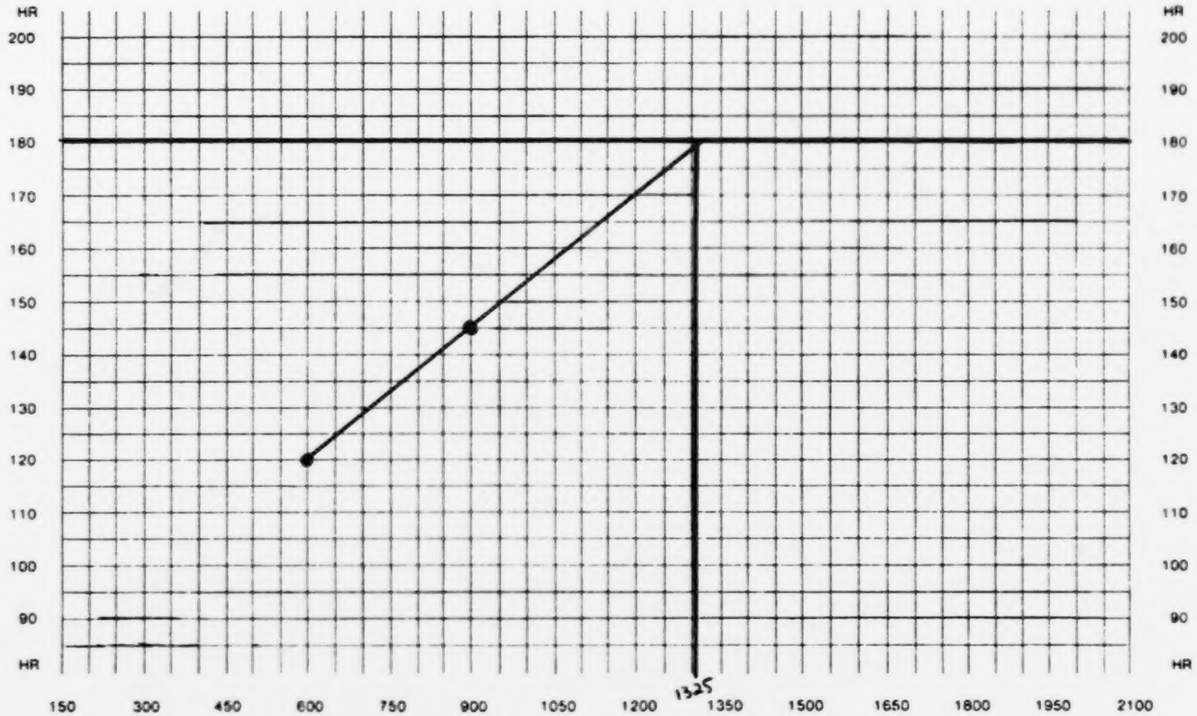
NAME Example Male AGE 40 WEIGHT 176 LB 80 KG SEAT HEIGHT 8

PREDICTED MAX HR 180

	DATE	1st WORKLOAD HR USED	2nd WORKLOAD HR USED	MAX WORKLOAD
TEST 1	1-4-82	600/120	900/145	1325
TEST 2				
TEST 3				

DIRECTIONS

1. Plot the HR of the 2 workloads versus the work (kgm/min).
2. Determine the subject's max HR line by subtracting subject's age from 220 and draw a line across the graph at this value.
3. Draw a line through both points and extend to the max HR line for age.
4. Drop a line from this point to the baseline and read the predicted max workload and 0_2 uptake.



WORKLOAD (kgm/min)	150	300	450	600	750	900	1050	1200	1350	1500	1650	1800	1950	2100
MAX 0_2 UPTAKE (L/m)	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.8	3.2	3.5	3.8	4.2	4.6	5.0
KCAL USED (Kcal/m)	3.0	4.5	6.0	7.5	9.0	10.5	12.0	14.0	16.0	17.5	19.0	21.0	23.0	25.0
APPROX. MET LEVEL (for 132) lbs.)	3.3	4.7	6.0	7.3	8.7	10.0	11.3	12.7	14.0	15.3	16.7	18.0	19.3	20.7
APPROX. MET LEVEL (for 176 lbs.)	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	11.0	12.0	13.0	14.0	15.0	16.0

Bicycle Ergometer Heart Rate Form
BICYCLE ERGOMETER
Physical Work Capacity Test

Age ________ Weight ________ lbs. ________ kg.

Seat Height ________ Predicted Max Heart Rate ________ B/M

WORKLOADS		**HEART RATE**	
1st Workload 300 kgm		________	2nd min
		________	3rd min
		________	4th min (if needed)
2nd Workload	________ kgm	________	2nd min
		________	3rd min
		________	4th min
3rd Workload	________ kgm	________	2nd min
		________	3rd min
		________	4th min (if needed)

Reprinted from *The Y's Way to Physical Fitness*, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

li

Y's WAY TO PHYSICAL

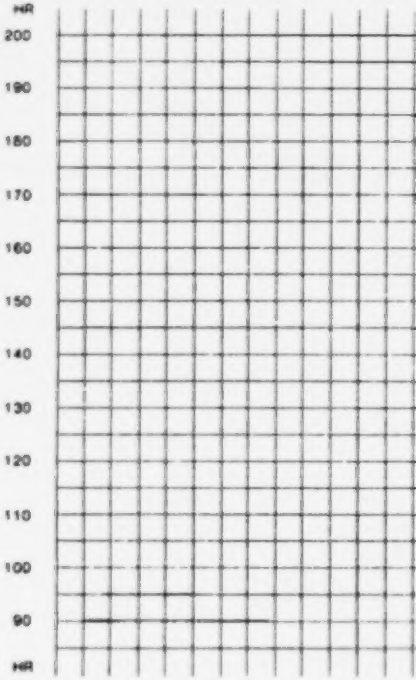
MAXIMUM PHYSICAL WO

NAME ______________________________ AGE____

	DATE	1st WORKLOAD HR USED	2nd WORKLOAD USED
TEST 1	________	________	________
TEST 2	________	________	________
TEST 3	________	________	________

DIRECTIONS

1. Plot the HR of the 2 workloads versus the work (kgm/min).
2. Determine the subject's max HR line by subtracting subject's age from 220 and draw a line across the graph at this value.
3. Draw a line through both points and extend to the max HR line for age.
4. Drop a line from this point to the baseline and read the predicted max workload and O_2 uptake.



WORKLOAD (kgm/min)	150	300	450	600	750
MAX O_2 UPTAKE (L/m)	0.6	0.9	1.2	1.5	1.8
KCAL USED (Kcal/m)	3.0	4.5	6.0	7.5	9.0
APPROX. MET LEVEL (for 132) lbs.)	3.3	4.7	6.0	7.3	8.7
APPROX. MET LEVEL (for 176 lbs.)	3.0	4.0	5.0	6.0	7.0

ction Graph

FITNESS – TEST BATTERY

RKING CAPACITY PREDICTION

____ WEIGHT ____________ LB __________ KG SEAT HEIGHT ____________

PREDICTED MAX. HR __________

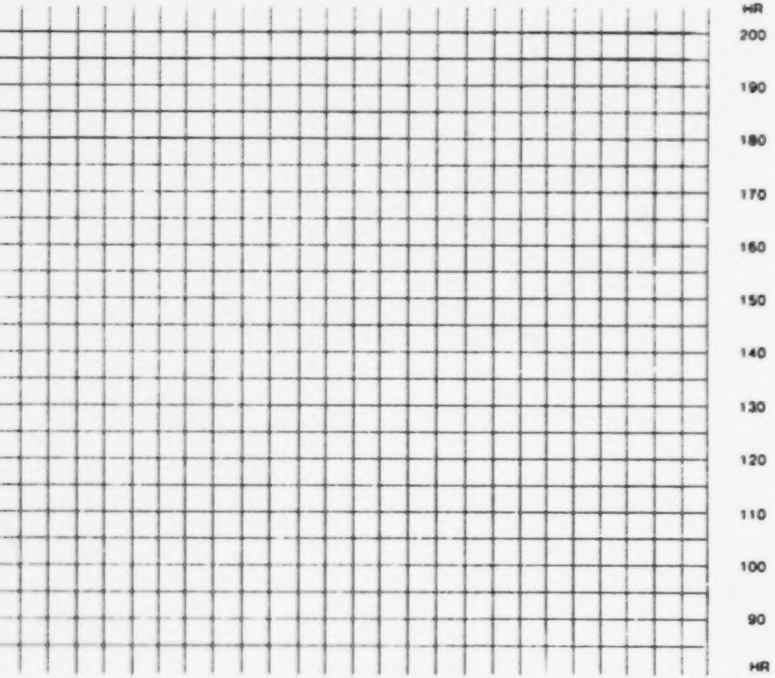
HR

MAX. WORKLOAD

____ ____________

____ ____________

____ ____________



900	1050	1200	1350	1500	1650	1800	1950	2100
2.1	2.4	2.8	3.2	3.5	3.8	4.2	4.6	5.0
10.5	12.0	14.0	16.0	17.5	19.0	21.0	23.0	25.0
10.0	11.3	12.7	14.0	15.3	16.7	18.0	19.3	20.7
8.0	9.0	10.0	11.0	12.0	13.0	14.0	15.0	16.0

3-MINUTE STEP TEST

This measure assesses participants' cardiorespiratory endurance. This measure is appropriate for adults.

PURPOSE

The 3-Minute Step Test can be successfully used in mass testing. However, it is also appropriate for self-administered, individual testing. Information regarding participants' cardiorespiratory endurance may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show that participants have a low level of aerobic capacity, thus indicating the need for a cardiorespiratory training program.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' cardiorespiratory endurance.

EQUIPMENT

- A sturdy bench, 12 inches high
- A metronome set at 96 bpm, or 24 steps per minute. (Four clicks of the metronome equals one step — up, up, down, down.)
- A timing clock
- A recovery clock (may be same as timing clock)
- A stethoscope to count recovery rate

ADMINISTRATIVE SUGGESTIONS

- Demonstrate the stepping procedure and have the participants pace the steps and pick up the beat of the metronome.
- Explain to participants the importance of sitting down quickly at the end of the three minutes and resting quietly for one minute so that the tester can take a heart rate.
- To help participants get a feel for the rhythm and maintain it throughout the test, use a cadence such as Step-Step UP, Step-Step DOWN or UP-2-3, DOWN-2-3.
- During the test be sure to check the rhythm and correct it if necessary. Also, call out the time as it passes, for example, "one minute, two minutes," etc.

Reprinted from the Y's Way to Physical Fitness with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

- Electrical and mechanical metronomes are available. Electrical metronomes have both an auditory and a visual signal and do not need to be wound. Mechanical metronomes are typically "music" metronomes with a wand that oscillates back and forth. The cadence is changed by moving a weight up or down the wand. Mechanical metronomes need to be wound. Both types of metronomes need to be calibrated. This is done by timing the number of beats with a stopwatch.
- Metronomes have no volume control and are often too quiet for testing. One answer to this problem is to make a tape recording of the metronome and use the recording during sessions. This will also eliminate the need for future calibration.

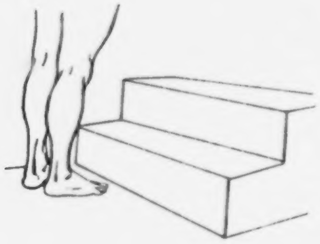
PROCEDURE

After explaining the test and allowing the participant to pick up the beat of the metronome, have the participant face the bench, and, in time with the metronome, step one foot up on the bench (1st beat), step up with the second foot (2nd beat), step down with the first foot (3rd beat), and step down with the other foot (4th beat). Begin the test and start keeping time. Remember to call out the time as it passes. When 20 seconds are left, remind participants to sit down quickly at the end of the stepping sequence and wait for the tester to take a heart rate. Prepare a recovery timer. On the last step it is helpful to say "Last step — up, up, down and sit down."

When the participant sits down, immediately place the stethoscope on the chest, get the rhythm, and start counting for one full minute. If participants are self-testing they should sit down and immediately count the carotid or radial pulse for one full minute. The recovery rate must be started within five seconds or the heart rate will be significantly different. The one-minute count reflects the heart's rate at the end of the test and also reflects the rate of recovery.

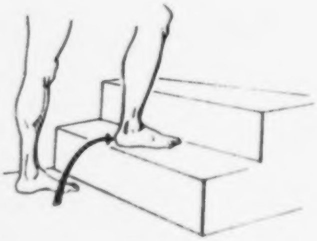
START:

Stand in front of the first step, feet together.



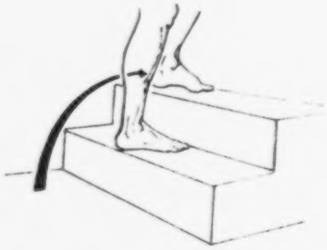
1. STEP:

Place your right foot up on the first step.



2. STEP:

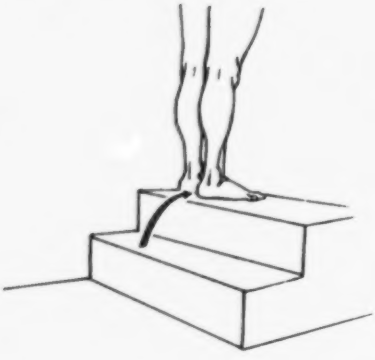
Bring your left foot up to the second step.



Reprinted from the Y's Way to Physical Fitness with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

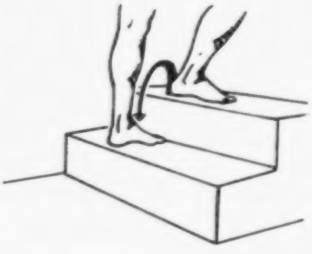
3. UP:

Bring your right foot up on the second step, feet together.



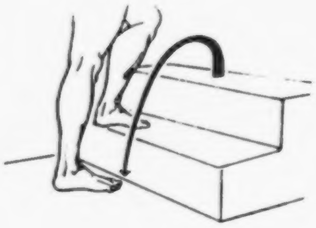
4. STEP:

Start down with your left foot to the first step.



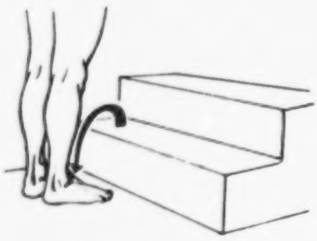
5. STEP:

Bring your right foot down to the ground level.



6. DOWN:

Bring your left foot to the ground level, feet together.



SCORING AND ANALYSIS

The total one-minute post-exercise heart rate is the score for the test. Score the total one-minute post-exercise heart rate, in beats per minute (bpm). Refer to the norm tables for percentage rankings. Those participants scoring below the 50th percentile should be encouraged to improve cardiovascular fitness. Those scoring below the 25th percentile should be encouraged to start a remedial cardiovascular program.

Drawings courtesy of Canada *Standardized Test of Fitness.*

TECHNICAL INFORMATION

In a series of five test-retest studies on the step test with two to seven days between tests, a mean difference of 1.46 ml/min per kg (maximum O_2 intake) following the retest was found (Kasch, 1966). The greatest difference was 1.6 ml/min per kg (maximum O_2 intake), which is well within the reproducibility limits of 2.1 ml/min per kg. suggested by Taylor et al. (Kasch, 1966; Taylor, 1955). The correlation between the bicycle ergometer and the step test is .95 (Kasch, 1966).

NORMS AND STANDARDS

3 MINUTE STEP TEST

One minute post-exercise heart rate

(in beats per minute)

Males

Percentage Ranking	Age		
	18-35	36-45	46+
95	81	84	90
85	99	98	102
75	103	112	111
50	120	120	120
30	123	125	124
15	127	129	130
5	136	138	138

Females

Percentage Ranking	Age		
	18-35	36-45	46+
95	79	79	84
85	94	90	97
75	109	106	108
50	118	118	118
30	122	125	124
15	129	134	130
5	137	145	145

Reprinted from the Y's Way to Physical Fitness with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

COOPER'S 1.5-MILE OR 12-MINUTE RUN/WALK

This measure is designed to estimate aerobic capacity or oxygen consumption. This measure is appropriate for adults and adolescents.

PURPOSE

Field testing of physical fitness is no longer a required part of Cooper's aerobics program, and is, in fact, contraindicated initially in the deconditioned person over 35 years of age. However, the 1.5-mile or 12-minute run/walk is an easy way to measure the success of a program and continues to be a popular feature of the aerobics system. It does not require expensive laboratory equipment and allows for large groups to be tested at one time. Information regarding participants' aerobic capacity may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show that participants have a low level of aerobic capacity, thus indicating the need for an aerobic training program.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' aerobic capacity.

EQUIPMENT

- A stopwatch or clock with a second hand
- An indoor or outdoor track or another suitable running area that can be measured

ADMINISTRATIVE SUGGESTIONS

- This measure is designed only for individuals who are either under 35 years of age, already conditioned, or have completed at least the first six weeks of an aerobic program.
- On the day of testing, participants should refrain from any physical exertion and should abstain for smoking or eating for two hours prior to the test.
- If possible, participants should practice pacing prior to the test. This helps participants not to run too fast early in the test. Another alternative is to have a trained pacer accompany participants during the test.
- Allow adequate time for warm-up exercises before the test and cool-down exercises after the test.
- It is recommended that participants work in pairs. This helps in the counting of laps and keeping track of times. If participants run at the same time, it is recommended that their times be called out as they finish.

PROCEDURE

1.5-Mile Run/Walk. Instruct participants to run or walk at maximum ability on a level surface for 1.5 miles. Start the stopwatch when participants start the test. As the participants finish, call out the time it took to run or walk the 1.5 miles.

12-Minute Run/Walk. Instruct participants to run or walk at maximum ability for 12 minutes. Stop participants after 12 minutes and measure the total distance covered. This is more easily accomplished by premeasuring the testing surface, such as a track, and having the partners count the number of laps completed.

SCORING AND ANALYSIS

The time or distance completed is the score. Participants that score below the category labeled "good" according to the Norms and Standards Table should be encouraged to improve cardiovascular endurance by means of a progressive training program.

TECHNICAL INFORMATION

Because the 1.5-mile or 12-minute run/walk is related to the maximum oxygen intake, it is considered a valid test of cardiorespiratory function and performance as well as an index of the participant's ability to run distances. Validity and test-retest correlation coefficients are .90 and .94, respectively (AAHPERD, 1980; Cooper, 1977).

NORMS AND STANDARDS

12-Minute Walking/Running Test†

Distance (Miles) Covered in 12 Minutes

		Age (years)					
Fitness Category		*13-19*	*20-29*	*30-39*	*40-49*	*50-59*	*60 +*
I. Very Poor	(men)	<1.30*	<1.22	<1.18	<1.14	<1.03	< .87
	(women)	<1.0	< .96	< .94	< .88	< .84	< .78
II. Poor	(men)	1.30-1.37	1.22-1.31	1.18-1.30	1.14-1.24	1.03-1.16	.87-1.02
	(women)	1.00-1.18	.96-1.11	.95-1.05	.88- .98	.84- .93	.78- .86
III. Fair	(men)	1.38-1.56	1.32-1.49	1.31-1.45	1.25-1.39	1.17-1.30	1.03-1.20
	(women)	1.19-1.29	1.12-1.22	1.06-1.18	.99-1.11	.94-1.05	.87-. 98
IV. Good	(men)	1.57-1.72	1.50-1.64	1.46-1.56	1.40-1.53	1.31-1.44	1.21-1.32
	(women)	1.30-1.43	1.23-1.34	1.19-1.29	1.12-1.24	1.06-1.18	.99-1.09
V. Excellent	(men)	1.73-1.86	1.65-1.76	1.57-1.69	1.54-1.65	1.45-1.58	1.33-1.55
	(women)	1.44-1.51	1.35-1.45	1.30-1.39	1.25-1.34	1.19-1.30	1.10-1.18
VI. Superior	(men)	>1.87	>1.77	>1.70	>1.66	>1.59	>1.56
	(women)	>1.52	>1.46	>1.40	>1.35	>1.31	>1.19

**<Means "less than"; > means "more than."*

---

†From *The Aerobics Way* by Kenneth H. Cooper, M.D., M.P.H. Copyright ©1977 by Kenneth H. Cooper. Reprinted by permission of Bantam Books. All rights reserved.

1.5-Mile Run Test†

Time (Minutes)

		Age (years)					
Fitness Category		*13-19*	*20-29*	*30-39*	*40-49*	*50-59*	*60 +*
I. Very Poor	(men)	>15:31*	>16:01	>16:31	>17:31	>19:01	>20:01
	(women)	>18:31	>19:01	>19:31	>20:01	>20:31	>21:01
II. Poor	(men)	12:11-15:30	14:01-16:00	14:44-16:30	15:36-17:30	17:01-19:00	19:01-20:00
	(men)	16:55-18:30	18:31-19:00	19:01-19:30	19:31-20:00	20:01-20:30	21:00-21:31
III. Fair	(men)	10:49-12:10	12:01-14:00	12:31-14:45	13:01-15:35	14:31-17:00	16:16-19:00
	(women)	14:31-16:54	15:55-18:30	16:31-19:00	17:31-19:30	19:01-20:00	19:31-20:30
IV. Good	(men)	9:41-10:48	10:46-12:00	11:01-12:30	11:31-13:00	12:31-14:30	14:00-16:15
	(women)	12:30-14:30	13:31-15:54	14:31-16:30	15:56-17:30	16:31-19:00	17:31-19:30
V. Excellent	(men)	8:37- 9:40	9:45-10:45	10:00-11:00	10:30-11:30	11:00-12:30	11:15-13:59
	(women)	11:50-12:29	12:30-13:30	13:00-14:30	13:45-15:55	14:30-16:30	16:30-17:30
VI. Superior	(men)	< 8:37	< 9:45	<10:00	<10:30	<11:00	<11:15
	(women)	<11:50	<12:30	<13:00	<13:45	<14:30	<16:30

<Means "less than"; > means "more than."

---

†From *The Aerobics Way* by Kenneth H. Cooper, M.D., M.P.H. Copyright ©1977 by Kenneth H. Cooper. Reprinted by permission of Bantam Books. All rights reserved.

DISTANCE RUN

This measure assesses the maximum functional capacity and endurance of the cardiorespiratory system. This measure is appropriate for preadolescents.

PURPOSE

Information about participants' cardiorespiratory endurance may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show that participants have a low level of aerobic capacity and endurance, thus indicating the need to participate in a cardiovascular training program.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' aerobic capacity and endurance.

EQUIPMENT

A track or any other flat measured area is needed. Examples of appropriately measured areas are the 440-yard or 400-meter track, 110-yard or 100-meter straightaway, other outside fields, or an indoor court area.

ADMINISTRATIVE SUGGESTIONS

- The distance run is not recommended for children with known medical problems that would be exacerbated by vigorous exercise.
- A proper warm-up and cool-down period should be included as part of the testing session.
- To obtain valid and reliable results, participants should be adequately prepared for the test. Proper preparation should include practicing distance running with an emphasis placed on pacing, as well as instruction on proper running, walking, and breathing techniques.
- Walking is permitted and participants should be informed that they can walk. However, participants should be encouraged to try to maintain a consistent pace and to walk for only short periods of time when necessary.
- Motivation is very important in obtaining good results. To help motivate children, fully explain the purpose of the test.
- It is recommended that participants work in pairs to help count laps and record times.

PROCEDURE

Both the one-mile run and the nine-minute run are designed for testing cardiorespiratory endurance. Standards and norms are provided for both. The decision as to which test to use may be based on such considerations as the availability of facilities and equipment, time limitations, administrative considerations, and the personal preference of the tester.

One-Mile Run. Participants are instructed to run one mile in the fastest time possible. As they cross the finish line, their times should be called out. Walking is permitted, but the objective is to cover the distance in the shortest possible time.

Nine-Minute Run. Participants are instructed to run as far as possible in nine minutes. Walking is permitted, but the objective is to cover as much distance as possible during the nine minutes.

SCORING AND ANALYSIS

The score consists of either the participant's time or distance covered. Participants can be compared with other persons of the same age and sex based on either the norms provided here or locally developed norms. Refer to the tables provided on the next page to determine percentile ranking. It should be noted that the results of a running test are *not* entirely determined by cardiorespiratory function. Genetic potential, body composition, efficiency, effort, and maturity also contribute to the test results. Thus, the results not only reflect cardiorespiratory fitness, but may also reflect inherited characteristics, running skill, relative leanness, and motivation to do well.

Participants who score below the 50th percentile should be encouraged to try to perform up to the median score as a minimum level of cardiorespiratory function. Participants who score above the 50th percentile should be encouraged to improve or maintain a score at or above approximately the 75th percentile.

TECHNICAL INFORMATION

Because the one-mile run and the nine-minute distance run are related to maximum oxygen intake, it is considered a valid test of cardiorespiratory function and performance as well as an index of the participant's ability to run distances. Distance runs have acceptable reliability when administered carefully and with properly prepared participants (AAHPERD, 1980).

Performance on distance runs of one mile or more have been shown to correlate significantly with maximal aerobic power. Correlation coefficients have varied between 0.22 to .90. The reliability of the distance run ranged from .75 to over .90. (AAHPERD, 1984).

NORMS AND STANDARDS

One-Mile Run - (Min/Seconds) - BOYS†

Percentile	Age							
	5	6	7	8	9	10	11	12
99	7:45	8:15	7:17	6:14	6:43	6:25	6:04	5:40
75	11:32	10:55	9:37	9:14	8:36	8:10	8:00	7:24
50	13:46	12:29	11:25	11:00	9:56	9:19	9:06	8:20
25	16:05	15:10	14:02	13:29	12:00	11:05	11:31	10:00
5	18:25	17:38	17:17	16:19	15:44	14:28	15:25	13:41

One-Mile Run - (Min/Seconds) - GIRLS†

Percentile	Age							
	5	6	7	8	9	10	11	12
99	9:03	8:06	7:58	7:45	7:21	7:09	7:07	6:57
75	13:09	11:24	10:55	10:35	9:58	9:30	9:12	8:36
50	15:08	13:48	12:30	12:00	11:12	11:06	10:27	9:47
25	17:59	15:27	14:30	14:16	13:18	12:54	12:10	11:35
5	19:00	18:50	17:44	16:58	16:42	17:00	16:56	14:46

9-Minute Run (Yards) - BOYS†

Percentile	Age							
	5	6	7	8	9	10	11	12
99	1975	2000	2400	2520	2450	2520	2520	2880
75	1320	1469	1683	1810	1835	1910	1925	1975
50	1170	1280	1440	1595	1660	1690	1725	1760
25	900	1090	1243	1380	1440	1487	1540	1500
5	600	816	990	1053	1104	1110	1170	1000

9-Minute Run (Yards) - GIRLS†

Percentile	Age							
	5	6	7	8	9	10	11	12
99	1584	1980	2340	2300	2450	2240	2170	2370
75	1300	1440	1540	1650	1835	1650	1723	1760
50	1140	1208	1358	1425	1660	1460	1480	1590
25	950	1017	1225	1243	1440	1250	1345	1356
5	700	750	970	960	1104	940	904	1000

†Reprinted by permission of the American Alliance for Health, Physical Education, Recreation and Dance, 1900 Association Drive, Reston, Virginia 22091

JACKSON AND POLLOCK'S SKIN FOLD MEASURES

This measure evaluates the level of subcutaneous adipose tissue (body fat) in relation to body composition. This measure is appropriate for adults.

PURPOSE

It is well documented that when individuals gain fat most of it occurs in subcutaneous areas in certain parts of the body. Using the thumb and forefinger, one can pinch this subcutaneous fat into skinfolds. As individuals get fatter these skinfolds get larger. Information about participants' skinfold thicknesses may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show that participants have a high percentage of body fat, thus indicating a need for participant education in the positive effects of exercise.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' body fat and body composition.

EQUIPMENT

Although quality plastic calipers may be used, the Lange Caliper is recommended because of its recognized precision.

ADMINISTRATIVE SUGGESTIONS

- Skinfold measurements should be taken prior to any physical activity because sweat and increased blood flow to the skin make measuring more difficult.
- All measurements should be taken on the same side of the body.
- To yield standard, reliable results, a tester should practice using the calipers. A tester should practice on the same group of individuals until results become consistent (within one to two millimeters) or two different testers could measure the same individual.

PROCEDURE

Grasp fold of skin firmly between thumb and four fingers, then lift up. Pinch and lift the fold several times to insure that you are not measuring muscle.

While firmly holding the skinfold with thumb and fingers, place the contact surface of the calipers below the thumb and fingers with your other hand. The jaws of the calipers must be placed exactly on the skinfold site location. If an adjustment

---

Reprinted from *The Y's Way to Physical Fitness*, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

is needed, adjust the finger grip, not the fold. Release the grip on the calipers completely, allowing the spring to compress the fold. When the needle on the caliper dial stops moving, take the reading to the nearest half-millimeter. Remove the caliper before releasing the fold.

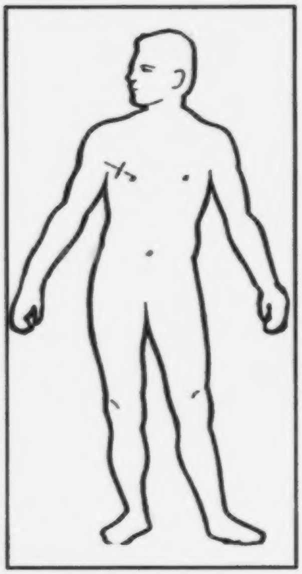
The following are the seven skinfold sites and descriptions of their locations:

1. **Chest** (pectoral): A diagonal fold on the pectoral line midway between the axillary fold and the nipple.
2. **Abdomen**: A vertical fold approximately one inch to the right of the umbilicus.
3. **Hip** (ilium or suprailium): A diagonal fold just above the crest of the ilium, that is, the highest peak on the side of the pelvic girdle on the mid-axillary line.
4. **Side** (axilla, mid-axilla): A vertical fold on the mid-axillary line at nipple level (mid-sternum).
5. **Arm** (triceps): A vertical fold on the back of the upper arm, midway between the shoulder and elbow joints.
6. **Back** (scapula, subscapula): A diagonal fold just below the inferior angle of the scapula.
7. **Thigh** (leg): A vertical fold on the front of the thigh, midway between groin line and the tip of the patella.

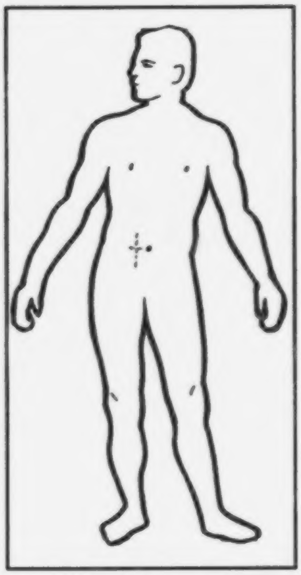
All seven sites should be measured to yield the most reliable results. If all seven sites cannot be measured, the following locations are acceptable alternatives:

women: triceps, abdomen (or thigh), suprailium
men: chest, axilla, abdomen (or thigh), suprailium

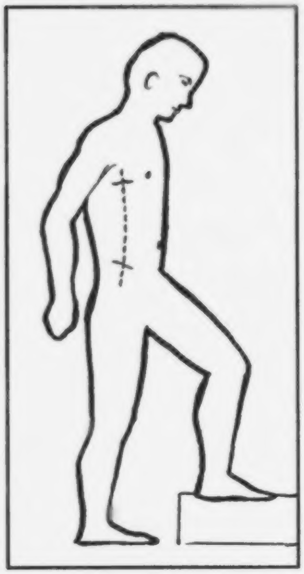
Reprinted from *The Y's Way to Physical Fitness*, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.



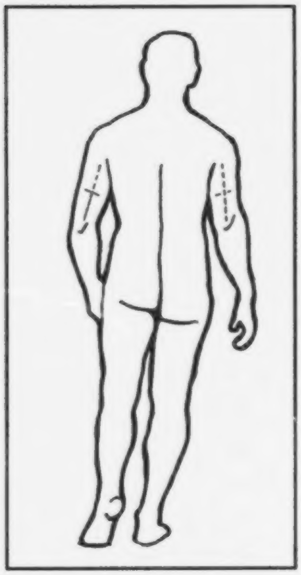
Chest (pectoral)



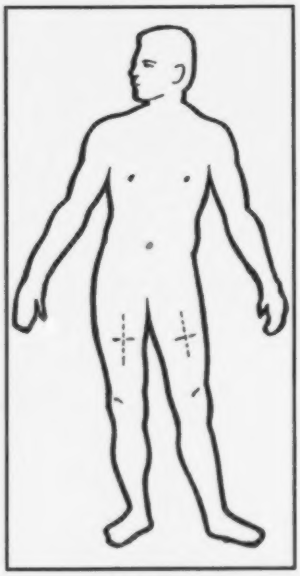
Abdomen



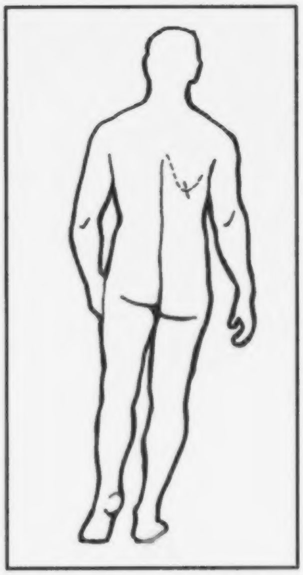
Hip (ilium) and Side (axilla)



Arm (triceps)



Thigh (leg)



Back (scapula)

Reprinted from *The Y's Way to Physical Fitness*, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

SCORING AND ANALYSIS

- Repeatedly measure the same location. To ensure consistency, each measurement should be within one to two millimeters. If repeated measurements are similar, record the last measurement taken to the nearest half-millimeter (.05). Refer to the tables in the norms and standards section for percentage ranking or percent fat estimates. The percentage ranking tables are for the individual scores for all seven sites. The percent fat estimates are the sum of the four alternative sites.

  Those participants scoring above the 50th percentile should be encouraged to achieve or maintain the degree of fatness between the 50th and 75th percentiles. Those scoring below the 50th percentile should be encouraged to reduce body fatness until their skinfold data reaches a more desired level (at least 50th percentile). Increased daily physical activity and reduced food intake are recommended behaviors for weight control and fat reduction.

TECHNICAL INFORMATION

The correlations (i.e., validity coefficients) between skinfolds and hydrostatically determined body fatness have consistently ranged from 0.70 to 0.90 in both children and adults. Hydrostatic weighing is an accepted and valid method used to measure the degree of body fatness. The test-retest reliability of skinfold fat measures has exceeded .95 in experienced testers (AAHPERD, 1980).

Reprinted from *The Y's Way to Physical Fitness*, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

BODY COMPOSITION RATING SCALE

Skinfolds

Males 18-35 Years Old

Percentage Ranking	Percent Fat	Chest mm	Abdomen mm	Ilium mm	Axilla mm	Tricep mm	Back mm	Thigh mm
95	6	3	4	4	4	3	4	4
85	9	7	8	6	8	6	8	6
75	14	12	16	11	13	10	12	10
50	18	15	21	16	17	11	15	14
30	22	18	27	20	21	13	19	16
15	25	22	34	26	25	16	24	21
5	30	28	44	33	33	21	33	33

Males 36-45 Years Old

Percentage Ranking	Percent Fat	Chest mm	Abdomen mm	Ilium mm	Axilla mm	Tricep mm	Back mm	Thigh mm
95	8	4	6	4	4	3	4	6
85	10	8	10	8	10	7	8	8
75	15	13	17	13	15	10	13	11
50	19	16	22	17	19	12	17	15
30	23	19	28	22	23	14	21	18
15	27	24	35	28	28	18	26	26
5	32	30	45	37	35	23	35	35

Males 46 Years and Older

Percentage Ranking	Percent Fat	Chest mm	Abdomen mm	Ilium mm	Axilla mm	Tricep mm	Back mm	Thigh mm
95	9	5	6	6	6	4	6	6
85	11	8	11	9	11	8	10	10
75	16	14	18	15	17	12	15	11
50	21	17	23	19	21	14	19	16
30	24	20	29	23	24	16	23	19
15	29	24	36	30	30	19	28	30
5	34	31	46	39	36	26	38	38

Reprinted from *The Y's Way to Physical Fitness*, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

BODY COMPOSITION RATING SCALE

Skinfolds

Females 18-35 Years Old

Percentage Ranking	Percent Fat	Chest mm	Abdomen mm	Ilium mm	Axilla mm	Tricep mm	Back mm	Thigh mm
95	9	5	5	4	6	5	4	5
85	14	9	8	7	11	7	6	8
75	18	15	14	13	17	12	8	17
50	22	18	19	16	19	15	12	24
30	24	26	25	20	24	19	16	30
15	28	34	33	29	30	25	22	39
5	35	40	40	35	36	30	27	42

Females 36-45 Years Old

Percentage Ranking	Percent Fat	Chest mm	Abdomen mm	Ilium mm	Axilla mm	Tricep mm	Back mm	Thigh mm
95	10	6	6	5	6	6	5	6
85	16	10	8	8	12	9	8	10
75	20	16	14	14	18	13	9	19
50	23	18	19	18	20	17	13	25
30	26	27	25	21	26	21	18	31
15	31	36	33	29	31	26	23	41
5	37	41	40	37	38	32	29	46

Females 46 Years and Older

Percentage Ranking	Percent Fat	Chest mm	Abdomen mm	Ilium mm	Axilla mm	Tricep mm	Back mm	Thigh mm
95	11	7	8	7	7	8	8	8
85	18	12	10	9	13	10	10	11
75	21	18	15	16	19	15	12	21
50	25	20	20	18	21	18	15	26
30	30	29	26	22	27	23	20	33
15	34	37	35	32	33	27	24	43
5	41	42	43	39	40	34	31	48

Reprinted from *The Y's Way to Physical Fitness*, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

PERCENT FAT ESTIMATES

Sum of Four Skinfolds
Chest, Ilium, Abdomen, Axilla

MALES

Sum of 4 Skinfolds	Age to Last Year								
	18 to 22	23 to 27	28 to 32	33 to 37	38 to 42	43 to 47	48 to 52	53 to 57	58 and older
8-12	1.9	2.5	3.2	3.8	4.4	5.0	5.7	6.3	6.9
3-17	3.3	3.9	4.5	5.1	5.7	6.4	7.0	7.6	8.2
18-22	4.5	5.2	5.8	6.4	7.0	7.7	8.3	8.9	9.5
23-27	5.8	6.4	7.1	7.7	8.3	8.9	9.5	10.2	10.8
28-32	7.1	7.7	8.3	8.9	9.5	10.2	10.8	11.4	12.0
33-37	8.3	8.9	9.5	10.1	10.8	11.4	12.0	12.6	13.2
38-42	9.5	10.1	10.7	11.3	11.9	12.6	13.2	13.8	14.4
43-47	10.6	11.3	11.9	12.5	13.1	13.7	14.4	15.0	15.6
48-52	11.8	12.4	13.0	13.6	14.2	14.9	15.5	16.1	16.7
53-57	12.9	13.5	14.1	14.7	15.4	16.0	16.6	17.2	17.9
58-62	14.0	14.6	15.2	15.8	16.4	17.1	17.7	18.3	18.9
63-67	15.0	15.6	16.3	16.9	17.5	18.1	18.8	19.4	20.0
68-72	16.1	16.7	17.3	17.9	18.5	19.2	19.8	20.4	21.0
73-77	17.1	17.7	18.3	18.9	19.5	20.2	20.8	21.4	22.0
78-82	18.0	18.7	19.3	19.9	20.5	21.0	21.8	22.4	23.0
83-87	19.0	19.6	20.2	20.8	21.5	22.1	22.7	23.3	24.0
88-92	19.9	20.5	21.2	21.8	22.4	23.0	23.6	24.3	24.9
93-97	20.8	21.4	22.1	22.7	23.3	23.9	24.5	25.2	25.8
98-102	21.7	22.3	22.9	23.5	24.2	24.8	25.4	26.0	26.7
103-107	22.5	23.2	23.8	24.4	25.0	25.6	26.3	26.9	27.5
108-112	23.4	24.0	24.6	25.2	25.8	26.5	27.1	27.7	28.3
113-117	24.1	24.8	25.4	26.0	26.6	27.3	27.9	28.5	29.1
118-122	24.9	25.5	26.2	26.8	27.4	28.0	28.6	29.3	29.9
123-127	25.7	26.3	26.9	27.5	28.1	28.8	29.4	30.0	30.6
128-132	26.4	27.0	27.6	28.2	28.8	29.5	30.1	30.7	31.3
133-137	27.1	27.7	28.3	28.9	29.5	30.2	30.8	31.4	32.0
138-142	27.7	28.3	29.0	29.6	30.2	30.8	31.4	32.1	32.7
143-147	28.3	29.0	29.6	30.2	30.8	31.5	32.1	32.7	33.3
148-152	29.0	29.6	30.2	30.8	31.4	32.1	32.7	33.3	33.9
153-157	29.5	30.2	30.8	31.4	32.0	32.7	33.3	33.9	34.5
158-162	30.1	30.7	31.3	31.9	32.6	33.2	33.8	34.4	35.1
163-167	30.6	31.2	31.9	32.5	33.1	33.7	34.3	35.0	35.6
168-172	31.1	31.7	32.4	33.0	33.6	34.2	34.8	35.5	36.1
173-177	31.6	32.2	32.8	33.5	34.1	34.7	35.3	35.9	36.6
178-182	32.0	32.7	33.3	33.9	34.5	35.2	35.8	36.4	37.0

Reprinted from *The Y's Way to Physical Fitness*, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

PERCENT FAT ESTIMATES

Sum of Three Skinfolds
Triceps, Abdomen, Ilium

FEMALE

Sum of 3 Skinfolds	Age to Last Year								
	18 to 22	23 to 27	28 to 32	33 to 37	38 to 42	43 to 47	48 to 52	53 to 57	58 and older
8-12	8.8	9.0	9.2	9.4	9.5	9.7	9.9	10.1	10.3
13-17	10.8	10.9	11.1	11.3	11.5	11.7	11.8	12.0	12.2
18-22	12.6	12.8	13.0	13.2	13.4	13.5	13.7	13.9	14.1
23-27	14.5	14.6	14.8	15.0	15.2	15.4	15.6	15.7	15.9
28-32	16.2	16.4	16.6	16.8	17.0	17.1	17.3	17.5	17.7
33-37	17.9	18.1	18.3	18.5	18.7	18.9	19.0	19.2	19.4
38-42	19.6	19.8	20.0	20.2	20.3	20.5	20.7	20.9	21.1
43-47	21.2	21.4	21.6	21.8	21.9	22.1	22.3	22.5	22.7
48-52	22.8	22.9	23.1	23.3	23.5	23.7	23.8	24.0	24.2
53-57	24.2	24.4	24.6	24.8	25.0	25.2	25.3	25.5	25.7
58-62	25.7	25.9	26.0	26.2	26.4	26.6	26.8	27.0	27.1
63-67	27.1	27.2	27.4	27.6	27.8	28.0	28.2	28.3	28.5
68-72	28.4	28.6	28.7	28.9	29.1	29.3	29.5	29.7	29.8
73-77	29.6	29.8	30.0	30.2	30.4	30.6	30.7	30.9	31.1
78-82	30.9	31.0	31.2	31.4	31.6	31.8	31.9	32.1	32.3
83-87	32.0	32.2	32.4	32.6	32.7	32.9	33.1	33.3	33.5
88-92	33.1	33.3	33.5	33.7	33.8	34.0	34.2	34.4	34.6
93-97	34.1	34.3	34.5	34.7	34.9	35.1	35.2	35.4	35.6
98-102	35.1	35.3	35.5	35.7	35.9	36.0	36.2	36.4	36.6
103-107	36.1	36.2	36.4	36.6	36.8	37.0	37.2	37.3	37.5
108-112	36.9	37.1	37.3	37.5	37.7	37.9	38.0	38.2	38.4
113-117	37.8	37.9	38.1	38.3	39.2	39.4	39.6	39.8	40.0
118-122	38.5	38.7	38.9	39.1	39.4	39.6	39.8	40.0	
123-127	39.2	39.4	39.6	39.8	40.0	40.1	40.3	40.5	40.7
128-132	39.9	40.1	40.2	40.4	40.6	40.8	41.0	41.2	41.3
133-137	40.5	40.7	40.8	41.0	41.2	41.4	41.6	41.7	41.9
138-142	41.0	41.2	41.4	41.6	41.7	41.9	42.1	42.3	42.5
143-147	41.5	41.7	41.9	42.0	42.2	42.4	42.6	42.8	43.0
148-152	41.9	42.1	42.3	42.8	42.6	42.8	43.0	43.2	43.4
153-157	42.3	42.5	42.6	52.8	43.0	43.2	43.4	43.6	43.7
158-162	42.6	42.8	42.0	43.1	43.3	43.5	43.7	43.9	44.1
163-167	42.9	43.0	43.2	43.4	43.6	43.8	44.0	44.1	44.3
168-172	43.1	43.2	43.4	43.6	43.8	44.0	44.2	44.3	44.5
173-177	43.2	43.4	43.6	43.8	43.9	44.1	44.3	44.5	44.7
178-182	43.3	43.5	43.7	43.8	44.0	44.2	44.4	44.6	44.8

Reprinted from *The Y's Way to Physical Fitness*, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

SUM OF SKINFOLD FAT

This measure evaluates the level of subcutaneous adipose tissue in the triceps and subscapula. These sites were selected because they are easily measured and are highly correlated with total body fat. This measure is appropriate for adolescents and preadolescents.

PURPOSE

Both coronary heart disease and stroke are determinants of atherosclerosis, a disease process known to begin in early childhood. Hypertension, hyperlipidemia, and physical inactivity are risk factors for atherosclerosis and are highly prevalent in children (Pate, 1983). Increasing evidence suggests that habitual physical activity is a key factor in improving body composition. Information about participants' body composition may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show that participants have above normal percentile scores, thus indicating the need for a maintenance program to sustain the positive scores.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' body composition.

EQUIPMENT

Although quality plastic calipers may be used, the Lange Caliper is recommended because of its recognized precision.

ADMINISTRATIVE SUGGESTIONS

- Participants need to wear a loose fitting T-shirts that may be pulled up in the back to measure the subscapular skin fold site. If a girl is wearing a bra, the strap needs only to be pushed upward two to three inches to allow for the measurement. Girls might want to wear a halter top or swimsuit top.
- Administer prior to any physical activity because sweat and increased blood flow to the skin makes measuring more difficult.
- Make sure the calipers are placed midway between the crest and base of the skinfold which is approximately one centimeter or slightly less than one half an inch below the fingers.
- To yield standard, reliable results, a tester should practice using the calipers. A tester should practice on the same group of individuals until results are consistent (within one to two millimeters) or, have two different testers measure the same individual, then check for consistency of the results.

- When measuring obese children, repeated tests may produce results with more than a two millimeter difference between scores. If this happens, it is recommended that an additional set of three measurements be taken. Record the average of the two middle scores.
- In all cases, scores should be measured separately for each child, without comment or display. Interpretation should also be given individually.
- Whenever possible, have the same person administer the test on the same persons at subsequent testing periods.

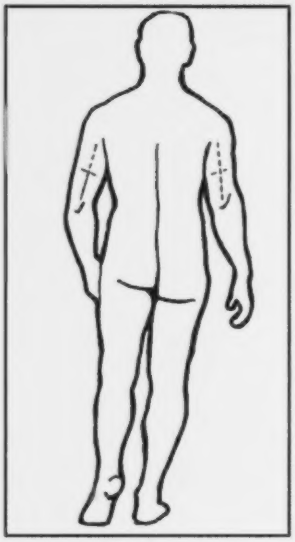
PROCEDURE

Grasp fold of skin firmly between thumb and four fingers, then lift up. Pinch and lift the fold several times to insure that you are not measuring muscle.

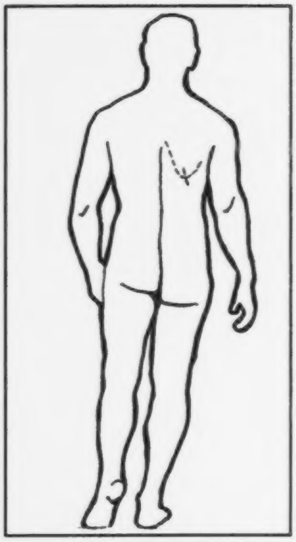
While firmly holding the skinfold with thumb and fingers, place the contact surface of the calipers below the thumb and fingers with your other hand. The jaws of the calipers must be placed exactly on the skinfold site location. If an adjustment is needed, adjust the finger grip, not the fold. Release the grip on the calipers completely, allowing the spring to compress the fold. When the needle on the caliper dial stops moving, take the reading to the nearest half-millimeter. Remove the caliper before releasing the fold.

The two skinfold sites and descriptions of their locations are as follows:

1. **Triceps:** A vertical fold on the back of the upper arm, halfway between the elbow and the acromion process of the scapula. The skinfold should be parallel to the longitudinal axis.
2. **Subscapula**: One centimeter (one-half inch) below the inferior angle of the scapula in line with the natural cleavage lines of the skin.



Triceps



Back (subscapula)

SCORING AND ANALYSIS

Add the measurements of the two sites and compare the scores to the norms included here. Those scoring above the 50th percentile are within the desired degree of body fat for children. Those scoring between the 50th and 25th percentiles should be encouraged to maintain the same weight level for the current year. Those children scoring below the 25th percentile should be given strong encouragement to reduce body fatness until their skinfold data reaches a more desired level. Increased daily physical activity and reduced food intake are recommended behaviors for weight control and fat reduction.

Skinfold measurements at the 90th percentile represent exceptional leanness. Participants that score above the 90th percentile should *not* be encouraged to lose weight. Reduction of weight at this level may result in loss of muscle and other non-fat tissues.

TECHNICAL INFORMATION

The correlations (i.e., validity coefficients) between skinfolds and hydrostatically determined body fatness have consistently ranged from .70 to .90 in both children and adults. Hydrostatic weighing is an accepted and valid method used by scientists to measure the degree of body fatness. The test-retest reliability of skinfold fat measures has exceeded .95 in experienced testers (AAHPERD, 1980).

NORMS AND STANDARDS

Percentile Norms
Ages 6-9 for the Sum of Triceps and Subscapular Skinfolds – BOYS (in millimeters)†

Percentile	Age			
	6	7	8	9
99	7	7	7	7
75	11	11	11	11
50	12	12	13	14
25	14	15	17	18
5	20	24	28	34

NCYFS Norms by Age for the Sum of Triceps and Subscapular Skinfolds – BOYS (in millimeters)††

Percentile	Age							
	10	11	12	13	14	15	16	17
99	9	9	9	9	9	10	10	10
95	11	11	10	11	11	11	11	12
90	12	11	11	12	12	12	12	13
80	13	13	13	13	13	13	14	14
75	13	14	14	14	14	14	14	14
70	14	15	14	14	14	14	14	15
60	15	16	15	15	16	15	16	16
50	17	17	17	17	17	17	17	18
40	19	19	19	19	18	18	19	19
30	22	23	21	21	21	20	21	21
25	23	24	23	22	22	22	23	22
20	25	27	25	25	24	24	25	24
10	36	39	32	34	31	31	32	28
5	44	47	40	45	42	37	40	35

---

†Reprinted by permission of the American Alliance for Health, Physical Education, Recreation and Dance, 1900 Association Drive, Reston, Virginia 22091.

††*National Children and Youth Fitness Study, Summary of Findings*, reprinted with permission of Glen Gilbert, Project Officer.

Percentile Norms Ages 6-9 for the Sum of Triceps and Subscapular Skinfolds – GIRLS (in millimeters)†

Percentile	Age			
	6	7	8	9
99	8	8	8	9
75	12	12	13	14
50	14	15	16	17
25	17	19	21	24
5	26	28	36	40

NCYFS Norms by Age for the Sum of Triceps and Subscapular Skinfolds – GIRLS†† (in millimeters)

Percentile	Age							
	10	11	12	13	14	15	16	17
99	11	11	11	12	13	14	13	15
95	12	13	14	14	15	17	17	18
90	14	15	15	17	17	19	19	20
80	15	16	18	19	20	22	21	22
75	16	17	18	20	21	23	22	23
70	17	18	19	20	22	24	23	24
60	19	20	21	22	25	25	25	26
50	21	22	24	25	27	28	27	28
40	23	25	26	27	29	30	30	30
30	26	28	28	30	32	32	32	33
25	28	29	30	32	34	34	34	35
20	31	32	33	34	36	37	36	36
10	37	40	41	40	43	42	42	42
5	46	50	49	47	49	52	49	47

---

†Reprinted by permission of the American Alliance for Health, Physical Education, Recreation and Dance, 1900 Association Drive, Reston, Virginia 22091.

††*National Children and Youth Fitness Study, Summary of Findings*, reprinted with permission of Glen Gilbert, Project Officer.

WAIST CIRCUMFERENCE

This measure is an index of deep adipose tissue. This measure is appropriate for adults.

PURPOSE

Information regarding participants' body circumference may be useful in the following ways:

- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' waist measurements.
- When used in conjunction with buttocks (hip) circumference, the waist-to-buttocks (hip) ratio is an indicator of the distribution of adipose tissue. The higher the waist-to-buttock (hip) ratio, the greater the risk of diseases such as noninsulin-dependent diabetes.
- Waist circumference is highly correlated with weight/stature, which is an index of general obesity.

EQUIPMENT

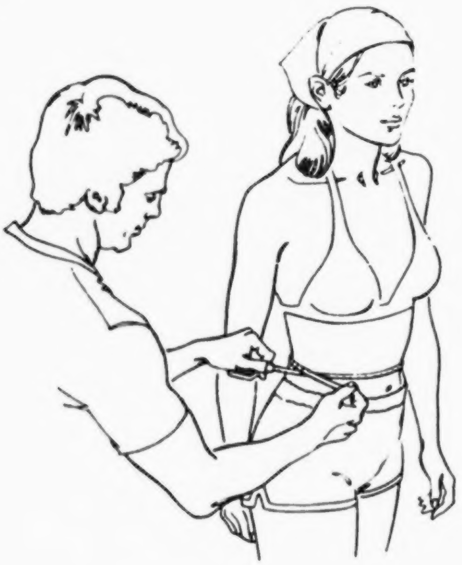
The only equipment needed is a flexible but inelastic (nonstretchable) tape measure which preferably has only one ruling on a side, (i.e., metric or English) and is about 0.7 cm wide.

ADMINISTRATIVE SUGGESTIONS

- Waist circumference is usually measured at the smallest circumference of the torso, which is at the level of the natural waist. Waist circumference measured at the level of the umbilicus will produce larger values.
- In obese participants it may be difficult to identify a natural waist. In such cases, the smallest horizontal circumference should be measured in the area between the ribs and iliac crest.
- Use of an assistant is very helpful in positioning and checking the tension of the measuring tape.

PROCEDURE

The participant stands erect with the abdomen relaxed, arms at sides and the feet together. The measurer faces the subject and places an inelastic tape around the participant, in a horizontal plane, at the level of the *natural* waist. Have an assistant check the horizontal plane of the tape and the tension. The measurement should be taken at the end of a normal expiration, without the tape compressing the skin.



SCORING AND ANALYSIS

The measurement should be recorded to the nearest 0.1 centimeter (cm). Norms for this measure do not take into consideration differences in body composition (percent fat/percent muscle) or different body proportions other than height and weight; thus, they are not included here. This measure should be used as an indicator for changes in body shape.

TECHNICAL INFORMATION

The technical error of measurement in adolescents in 1.31 cm for intrameasurer errors and 1.56 cm for intermeasurer errors (Malina, 1973). The technical error of measurement in the elderly is 0.48 cm in men and 1.15 cm in women (Chumlea, 1984). Thus, the "true" measurement of an individual would typically be within plus or minus 1 cm of the measured value in most cases.

Reprinted with permission of the publisher. *Anthropometric Standardization Reference Manual*, Timothy G. Lohman, Alex F. Roche, Reynaldo Martorell (Ed.). Human Kinetics Publishers, Inc.

Drawing courtesy of Canada *Standardized Test of Fitness*.

BUTTOCKS (HIP) CIRCUMFERENCE

This measure is an indicator of lower body fatness and reflects the amount of adipose tissue in the pelvic region. This measure is appropriate for adults.

PURPOSE

Information regarding participants' hip circumference may be useful in the following ways:

- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' body shape.
- When used in conjunction with waist circumference, the waist-to-hip circumference ratio is an indicator of the pattern of subcutaneous adipose tissue distribution. The higher the waist-to-hip ratio, the greater the risk of diseases such as Type II diabetes.

EQUIPMENT

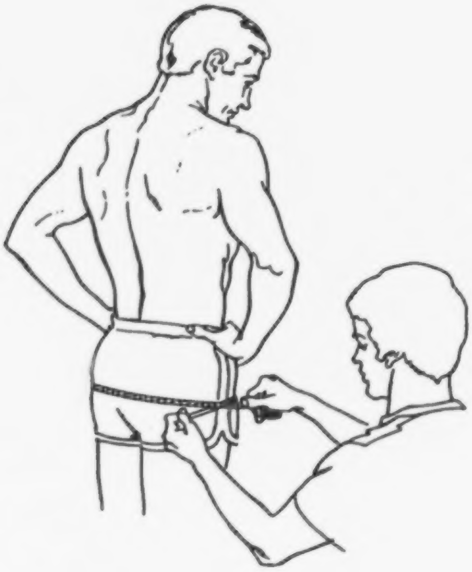
The only equipment needed is a flexible but inelastic (nonstretchable) tape measure that preferably has only one ruling on a side (i.e., metric or English) and is about 0.7 cm wide.

ADMINISTRATIVE SUGGESTIONS

- The buttocks (hip) circumference should be measured horizontally at the level of maximum extension of the posterior. This site is easiest to locate and yields the most accurate measurements.
- In very obese participants, the anterior abdominal wall may sag and, as a result, will be included in the measurement.

PROCEDURE

The participant should stand erect with arms at the sides and the feet together. The measurer should squat at the side of the participant so that the level of maximum extension of the buttocks can be seen. Place the tape around the buttocks horizontally without compressing the skin. Have an assistant check for horizontal placement and tension of the tape. The zero end of the tape should be below the measurement value.



SCORING AND ANALYSIS

The measurement should be recorded to the nearest 0.1 cm. Norms for this measure do not take into consideration differences in body composition (percent fat/percent muscle) or different body proportions other than height and weight; thus, they are not included here. This measure should be used as an indicator for changes in body shape.

TECHNICAL INFORMATION

Little is known about the reliability of hip circumference measurements. In a U.S. national survey of adolescents, the technical error was 1.23 cm for intrameasurer errors and 1.38 cm for intermeasurer errors (Malina, 1973). Therefore, in most cases, the "true" measurement for an individual would typically be within plus or minus 1 cm of the measurement recorded.

Reprinted with permission of the publisher. *Anthropometric Standardization Reference Manual*, Timothy G. Lohman, Alex F. Roche, Reynaldo Martorell (Ed.). Human Kinetics Publishers, Inc. Drawings courtesy of Canada *Standardized Test of Fitness*.

PUSH-UPS (CANADIAN FITNESS TEST)

This measure is designed to assess upper body (triceps, anterior deltoids, and pectoralis major) muscular strength and endurance. This measure is appropriate for adults.

PURPOSE

The upper body strength and endurance needed for push-ups contributes to injury-free performance of daily activities. Information on participants' upper body strength and endurance may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show that participants have a low level of upper body strength, thus indicating a need for training in that area.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' upper body strength and endurance.

EQUIPMENT

Exercise mat or other cushioned surface (e.g. flat, grassy area)

ADMINISTRATIVE SUGGESTIONS

- If a participant is over the age of 50 or in poor physical condition, use of this measure should be carefully considered.
- Push-ups are to be performed consecutively and without a time limit.
- The test should be stopped as soon as a participant strains forcibly to complete a push-up.
- Participants should work in pairs. The "helper" should prevent the feet of the participant being tested from slipping during administration and also help determine when the participant should stop testing.

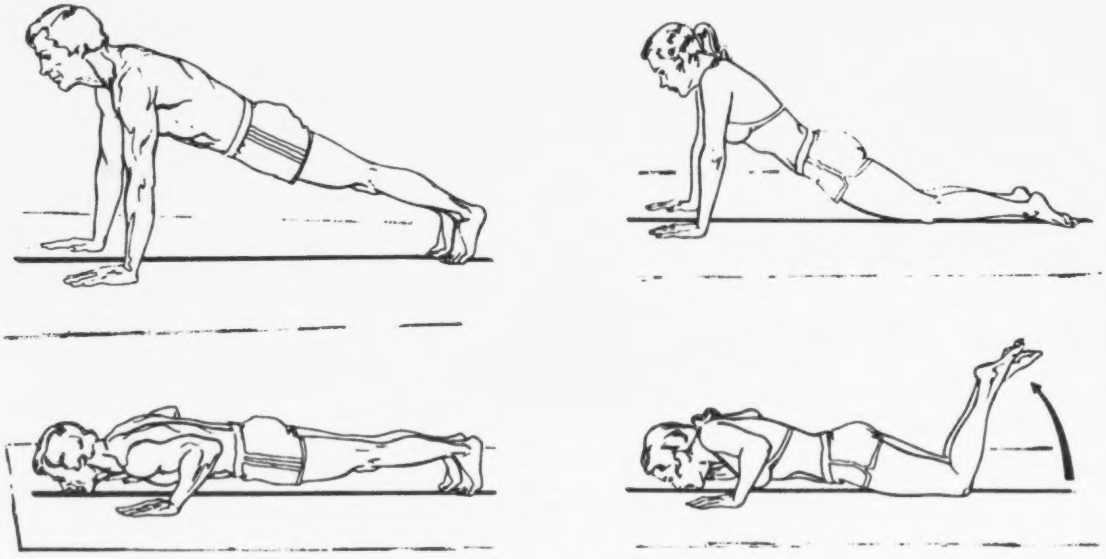
PROCEDURE

Separate procedures are provided for men and women.

Men. The participant lies on his stomach, legs together. The hands should be pointing forward, and positioned under the shoulders. The participant pushes up from the mat by straightening the elbows and using the *toes* as the pivotal point. The body must be kept in a straight line. The participant then returns to the starting position, *chin to the mat*.

Women. The participant lies on her stomach, legs together. The hands should be pointing forward and positioned under the shoulders. The participant pushes up from the mat by straightening the elbows and using the *knees* as the pivotal point.

The upper body must be kept in a straight line. As the participant returns to the starting position, *chin to the mat*, her feet should swing upward simultaneously.



SCORING AND ANALYSIS

Record the number of repetitions successfully completed. Participants scoring below the 50th percentile on the norms and standards tables should be encouraged to participate in a strength development program for the upper body.

TECHNICAL INFORMATION

Push-up, chin-up, and flexed-arm hang measures were developed to assess upper body strength and endurance. However, there does not appear to be any single isotonic test that offers a complete assessment of upper body strength (YMCA, 1982).

PUSH-UPS†

Males

Percentile	Age					
	17-19	20-29	30-39	40-49	50-59	60-65
95	49	41	35	29	26	25
85	41	35	29	25	22	20
75	36	31	26	22	19	17
50	27	23	19	16	13	12
30	21	18	14	12	9	7
15	14	12	9	7	5	3
5	6	6	4	3	1	0

Females

Percentile	Age					
	17-19	20-29	30-39	40-49	50-59	60-65
95	32	32	32	37	22	20
85	28	26	26	22	18	16
75	25	23	22	19	15	13
50	19	17	16	13	11	9
30	14	12	10	9	7	5
15	9	8	8	4	3	2
5	4	2	2	1	0	0

---

†Courtesy of Canada *Standardized Test of Fitness*

FITNESSGRAM® FLEXED-ARM HANG

This measure assesses upper body strength and endurance. This test is appropriate for adolescents and preadolescents.

PURPOSE

Sufficient upper body strength and endurance help reduce the stress on other body systems, such as the cardiovascular system, and are also required to maintain proper body alignment. Information regarding participants' upper body strength and endurance may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show that participants lack appropriate upper body strength, thus indicating a need for a training program in that area.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' upper body strength and endurance.

EQUIPMENT

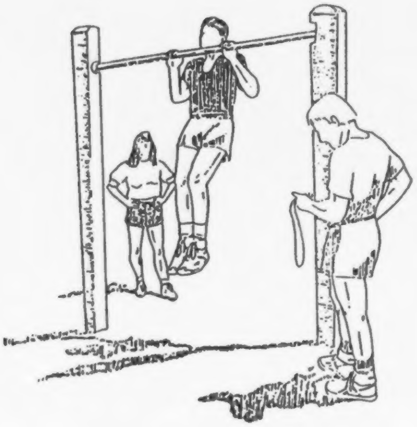
- A horizontal bar that is only slightly above a participant's standing height
- A stopwatch

ADMINISTRATIVE SUGGESTIONS

A participant's body should *not* swing during the test. If a participant starts to swing, the tester should place an extended arm across the front of the thighs to prevent the swinging motion.

PROCEDURE

The participant should use an overhand grasp (palms facing forward). With assistance of one or more spotters, the participant raises the body off the floor to a position where the chin is above the bar, the elbows are flexed, and the chest is close to, but not touching, the bar. Start the stopwatch as soon as the participant takes the hanging position. The hanging position is held as long as possible. The test stops when the participant's chin touches the bar, the head tilts backward to keep the chin above the bar, or the participant's chin falls below the level of the bar.



Flexed-arm hang

SCORING AND ANALYSIS

The score is the number of seconds the participant is able to maintain the correct hanging position. Refer to the norms and standards to determine the number of seconds required to achieve the minimum level of performance that is consistent with good health. A below average score is indicative of a low level of upper body strength. People who score below average should be encouraged to participate in a strength development program for the upper body. In some cases, a lack of upper body strength may be complicated by excess body fat. Improved body composition may also help improve the performance on this test.

TECHNICAL INFORMATION

Push-up, chin-up, and flexed-arm hang measures were developed to assess upper body strength and endurance. However, there does not appear to be any single isotonic test that offers a complete assessment of upper body strength (YMCA, 1982).

NORMS AND STANDARDS

FITNESSGRAM® compares the participant's performance to health standards that have been established by a panel of experts nationwide (Fitnessgram®, 1979). The health-referenced standard is considered the minimum level of performance on the flexed-arm hang that is consistent with good health.

Flexed-Arm Hang (seconds) – BOYS

Age	5	6	7	8	9	10	11	12	13	14	15	16+
Seconds	5	5	5	10	10	10	10	10	15	25	25	25

Flexed-Arm Hang (seconds) – GIRLS

Age	5	6	7	8	9	10	11	12	13	14	15	16+
Seconds	5	5	5	8	8	8	8	8	12	12	12	12

Fitnessgram® reprinted with permission of the Institute of Aerobics Research, Dallas, Texas.

SIT-UPS

This measure assesses the muscular strength and endurance of the abdominal muscles. This measure is appropriate for adults.

PURPOSE

Weak abdominal muscles are a contributing factor in the development of low-back pain and associated problems. Information on participants' abdominal muscle strength and endurance may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show that participants lack appropriate abdominal strength, thus indicating a need for training in that area.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' abdominal strength and endurance.

EQUIPMENT

- Exercise mat or other cushioned surface (e.g., flat, grassy area)
- Stopwatch or a watch with a second hand

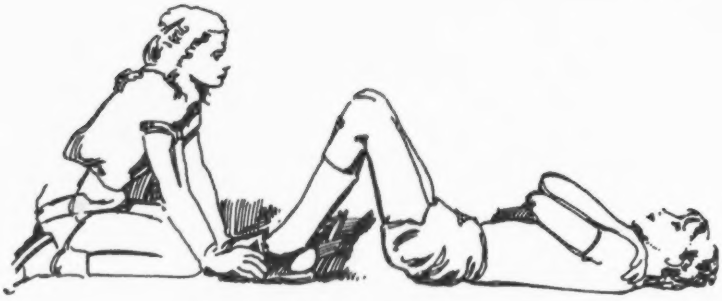
ADMINISTRATIVE SUGGESTIONS

- Prior to testing, participants should be instructed in appropriate sit-up technique and be allowed sufficient time to practice.
- Participants should work in pairs. The "helper" should hold the feet of the individual being tested and count the number of correctly executed sit-ups.
- To yield more valid and reliable results, only the tops of the feet should be held. Contact should not be made with the ankles or legs.
- Resting between sit-ups is accepted; however, it should be emphasized that there is a 60 second time limit and that participants are to perform as many correctly executed sit-ups as they can in that time period.
- The feet should be appropriately distanced from the buttocks and be in contact with the testing surface.

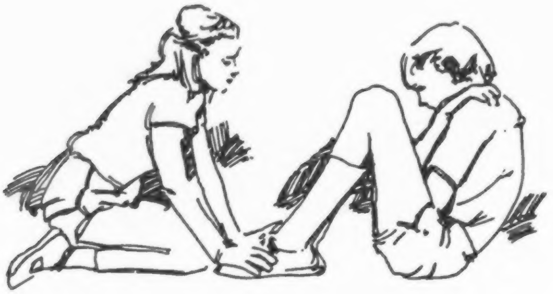
PROCEDURE

Participants lie on their backs with the knees flexed and feet on the floor. Both the feet and buttocks should be on the same surface. The arms are crossed on the chest with the hands on opposite shoulders. By tightening the abdominal muscles, participants curl into a sitting position. The chin should remain tucked on the chest.

After the elbows touch the thighs, participants return to the down position until the midback makes contact with the testing surface.



Starting position for the sit-up test



Up position for the sit-up test

SCORING AND ANALYSIS

Record the number of correctly executed sit-ups that are completed in 60 seconds. Refer to the norm table for the percentile scores. Participants who score below the 50th percentile should be encouraged to improve abdominal strength and endurance along with low back, hip, and thigh flexibility. For those scoring under the 25th percentile, a remedial program should be initiated.

TECHNICAL INFORMATION

Studies of muscle activity have shown that during the execution of a sit-up, abdominal muscles are utilized (AAHPERD, 1980). Sit-up tests are generally reliable; the test-retest reliability coefficients range from approximately .68 to .94 (AAHPERD, 1984).

NORMS AND STANDARDS

SIT-UPS†
(1 minute repetitions)

Males

Percentage Ranking	Age		
	18-35	36-45	46+
95	45	42	38
85	41	38	33
75	37	32	26
50	33	27	21
30	28	21	18
15	23	18	15
5	18	11	10

Females

Percentage Ranking	Age		
	18-35	36-45	46+
95	39	39	24
85	34	29	20
75	30	22	17
50	25	18	14
30	20	12	11
15	15	9	7
5	10	4	2

---

†Reprinted from *The Y's Way to Physical Fitness*, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

MODIFIED SIT-UPS

This measure assesses the strength and endurance of the abdominal muscles. This measure is appropriate for adolescents and preadolescents.

PURPOSE

Weak abdominal muscles are a contributing factor to the development of low-back pain and associated problems. Having information on participants' abdominal strength and endurance may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show that participants lack appropriate abdominal strength and endurance, thus indicating a need for a training program in that area.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' abdominal strength and endurance.

EQUIPMENT

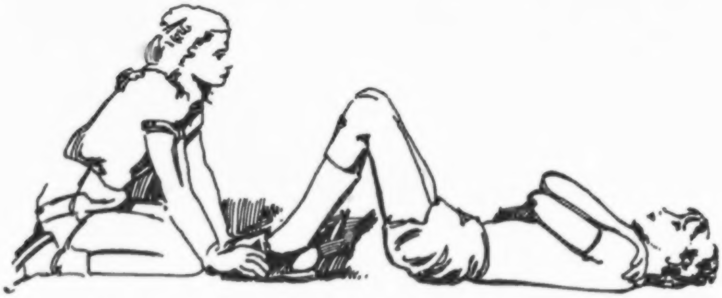
- Mats or other comfortable flat surfaces
- A stopwatch or a watch or clock with a second hand

ADMINISTRATIVE SUGGESTIONS

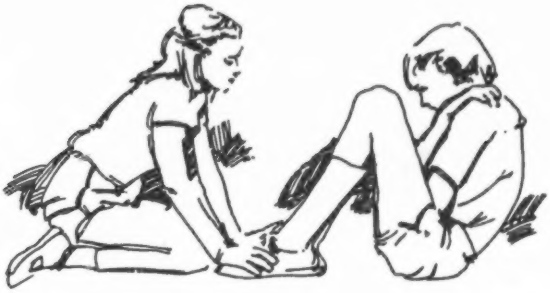
- Participants should work in pairs. One partner should hold the feet of the individual being tested and count the number of correctly performed sit-ups.
- The heels of the feet should remain the proper distance (12-18 inches) from the buttocks and be in contact with the testing surface.
- Participants should be reminded to keep buttocks on the floor to avoid a rocking motion.
- Prior to testing, participants should be instructed in appropriate sit-up technique, and be allowed sufficient time to practice.
- Participants should be informed that resting between sit-ups is permitted. However, it should be emphasized that there is a 60 second time limit and that participants are to perform as many correctly executed sit-ups as they can in that time period.

PROCEDURE

Participants lie on their backs with the knees flexed, feet on the floor, and with the heels 12-18 inches from the buttocks. Both the feet and buttocks should be on the same surface. The arms are crossed on the chest with the hands on opposite shoulders. By tightening the abdominal muscles, participants curl into a sitting position. The chin should remain tucked on the chest. After the elbows touch the thighs, participants return to the down position until the midback makes contact with the testing surface.



Starting position for the sit-up test



Up position for the sit-up test

SCORING AND ANALYSIS

Record the number of correctly executed sit-ups that are completed in 60 seconds. Refer to the norm table for the percentile scores. Participants who score below the 50th percentile should be encouraged to improve abdominal strength and endurance along with low back, hip, and thigh flexibility. For those scoring under the 25th percentile, a remedial program should be initiated.

TECHNICAL INFORMATION

Studies of muscle activity have shown that during the execution of a sit-up, abdominal muscles are utilized (AAHPERD, 1980). Sit-up tests are generally reliable; the test-retest reliability coefficients range from approximately .68 to .94 (AAHPERD, 1984).

NORMS AND STANDARDS

Modified Sit-Ups – BOYS†

Percentile	Age												
	5	6	7	8	9	10	11	12	13	14	15	16	17+
99	47	47	53	55	52	59	61	68	70	70	69	70	65
75	23	26	33	37	38	40	42	46	48	49	49	51	52
50	18	20	26	30	32	34	37	39	41	42	44	45	46
25	11	15	19	25	25	27	30	31	35	36	38	38	38
5	2	6	10	15	15	15	17	19	25	27	28	28	25

Modified Sit-Ups – GIRLS†

Percentile	Age												
	5	6	7	8	9	10	11	12	13	14	15	16	17+
99	35	42	51	55	51	54	55	61	60	57	64	63	65
75	24	28	31	35	35	39	40	41	41	42	43	42	44
50	19	22	25	29	29	32	34	36	35	35	37	33	37
25	12	14	20	22	23	25	28	30	29	30	30	29	31
5	2	6	10	12	14	15	19	19	18	20	20	20	19

---

†Reprinted by permission of the American Alliance for Health, Physical Education, Recreation and Dance, 1900 Association Drive, Reston, Virginia 22091.

SIT AND REACH TEST

This measure assesses the flexibility (extensibility) of the low back and posterior thighs. This measure is appropriate for adults.

PURPOSE

Lack of flexibility in the low back and posterior thighs is considered a precursor to low-back pain and other musculoskeletal problems. Information about participants' low-back and posterior-thigh flexibility may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show that participants have normal flexibility of the low back and posterior thighs, thus indicating a need for a flexibility maintenance program.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' low-back and posterior-thigh flexibility.

EQUIPMENT

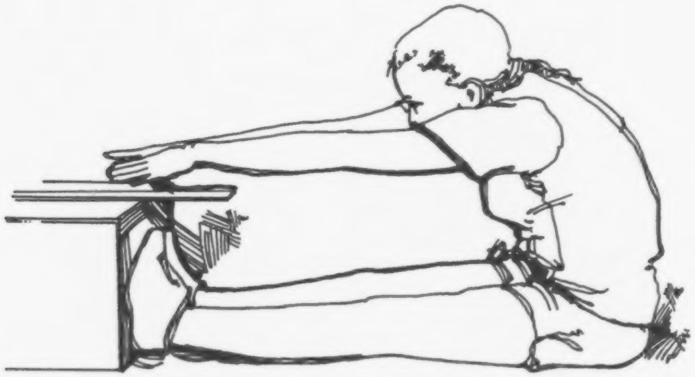
A specially constructed box with a measuring scale that has one centimeter (cm) gradations is needed. The 23 cm mark should be placed where the feet are located when the legs are fully extended. It is recommended that a box that has been constructed according to the suggested guidelines be used to ensure standardization of test procedures. However, if the box cannot be made, a bench with a metric ruler attached can be used. Regardless of the apparatus used, the feet should be at the 23 cm mark; otherwise, the norm tables will be invalid. Directions for constructing the Sit and Reach Box are provided.

ADMINISTRATIVE SUGGESTIONS

- The reliability and validity of the test can be improved by providing sufficient instruction and time for warm-up. The warm-up should include slow, sustained static stretching of the low back and posterior thighs.
- The test should be repeated if the participants' hands reach out unevenly or the knees are flexed. The flexing of knees can be prevented by having the tester place both hands lightly across the knees.
- Place the apparatus against a wall or another immovable object to prevent it from slipping away from the participants while testing.

PROCEDURE

Have participants remove their shoes and sit down at the test apparatus with their knees fully extended and the feet shoulder-width apart. The feet should be flat against the board. The arms are extended forward with the hands placed on top of each other to perform the test. Without bouncing or jerking, participants reach directly forward, palms down, along the measuring scale four times. The position of maximum reach on the fourth trial should be held for one second.



Test position for the sit and reach test

SCORING AND ANALYSIS

The score is the most distant point reached on the fourth trial measured to the nearest centimeter. This point must be touched by the fingertips of *both* hands. Refer to tables for percentile ranking. Scores below the 50th percentile represent poor extensibility in the lower trunk areas: posterior thigh, low back, or posterior hip.

The location of musculature tightness may be determined by observing participants as they perform the test. A poor performance that is coupled with a "rounding" of both the upper and lower back suggests tightness in the posterior thigh. Poor performance accompanied by a "rounding" of the upper back and a relatively "straight" lower back suggests a tightness in the lower back musculature.

RELIABILITY AND VALIDITY

The sit and reach test has been validated against several other types of flexibility tests. The coefficients obtained have generally ranged between .80 and .90. Reliability coefficients for the sit and reach test have been high, ranging above .70 (AAHPERD, 1980).

NORMS AND STANDARDS

SIT AND REACH†

Males

Percentile Ranking	Age					
	18-35		36-45		46 +	
	ins	cm	ins	cm	ins	cm
95	23	58	23	58	22	56
85	21	53	21	53	19	48
75	20	51	19	48	18	46
50	18	46	17	43	15	38
30	15	38	14	36	14	36
15	14	36	12	30.5	11	28
5	11	28	10	25	9	23

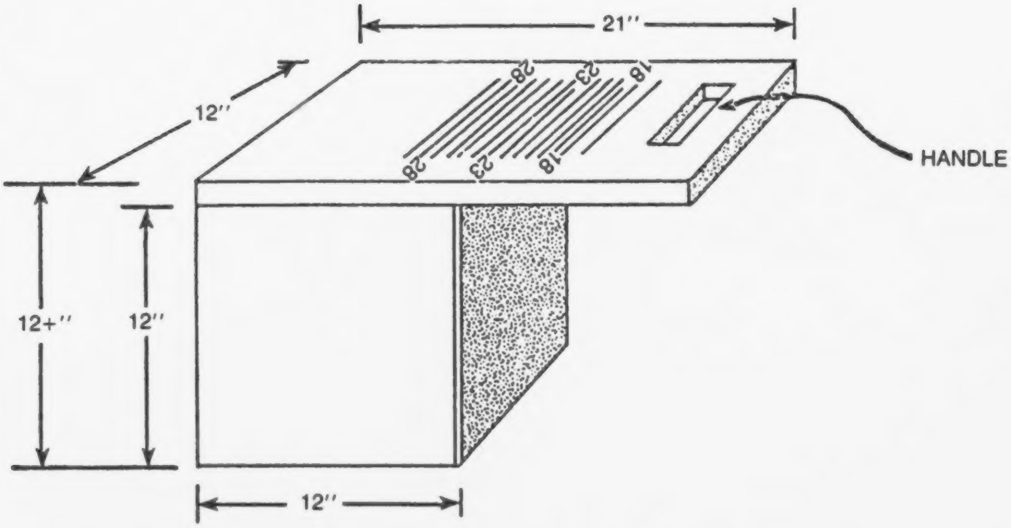
Females

Percentile Ranking	Age					
	18-35		36-45		46 +	
	ins	cm	ins	cm	ins	cm
95	21	53	22	56	20	51
85	19	48	19	48	17	43
75	17	43	16	41	15	38
50	15	38	14	36	13	33
30	12	30.5	12	30.5	11	28
15	9	23	10	25	8	20
5	7	18	5	13	5	13

---

†Reprinted from The Y's Way to Physical Fitness, with permission of the YMCA of the USA, 101 N. Wacker Drive, Chicago, Illinois 60606.

SIT AND REACH BOX



Schematic drawing of the sit and reach measurement apparatus

1. Using any sturdy wood or comparable construction material (3/4 inch plywood seems to work well), cut the following pieces:
    2 pieces - 12 in x 12 in
    2 pieces - 12 in x 10 1/2 in
    1 piece - 12 in x 21 in
2. Assemble the pieces using nails or screws and wood glue.
3. Inscribe the top panel with one centimeter gradations. It is crucial that the 23 centimeter line be exactly in line with the vertical panel against which the subject's feet will be placed.
4. Cover the apparatus with two coats of polyurethane sealer or shellac.
5. For convenience, a handle can be made by cutting a 1 inch x 3 inch hole in the top panel.
6. The measuring scale should extend from about 9 to about 50 cm.

AAHPERD SIT AND REACH TEST

This measure assesses the flexibility (extensibility) of the low back and posterior thighs. This test is appropriate for adolescents and preadolescents.

PURPOSE

Lack of flexibility in the low back and hamstring musculature has been identified as a precursor of low-back pain and other musculoskeletal problems. Information about participants' low-back and posterior-thigh flexibility may be useful in the following ways:

- Administration of this measure at the beginning of the program may provide needs assessment information. For example, results of this measure may show that participants lack flexibility in the low back and hamstring musculature, thus indicating the need for a flexibility training program.
- When this measure is administered prior to and following a program, it is possible to evaluate changes in participants' low-back and posterior-thigh flexibility.

EQUIPMENT

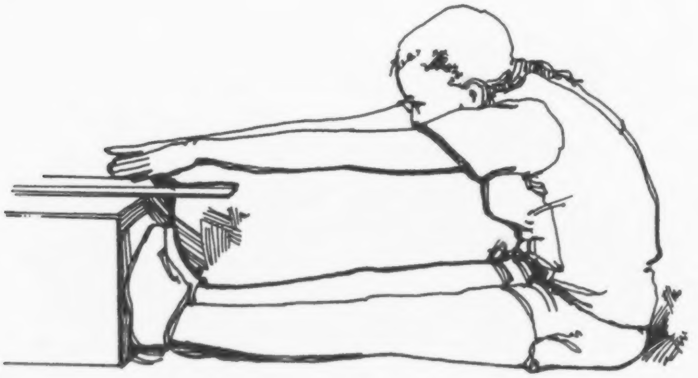
A specially constructed box with a measuring scale that has one centimeter (cm) gradations. The 23 cm mark should be placed where the feet are located when the legs are fully extended. It is recommended that a box that has been constructed according to the suggested guidelines be used to ensure standardization of test procedures. However, if the box cannot be made, a bench with a metric ruler attached can be used. Regardless of the apparatus used, the feet should be at the 23 cm mark; otherwise, the norm tables will be invalid. Directions for constructing the Sit and Reach Box are provided later.

ADMINISTRATIVE SUGGESTIONS

- The reliability and validity of the test can be improved by providing sufficient instruction and time for warm-up. The warm-up should include slow sustained static stretching of the low back and posterior thighs.
- The test should be repeated if the participants' hands reach out unevenly or the knees are flexed. The flexing of knees can be prevented by having the tester place both hands lightly across the knees.
- Place the apparatus against a wall or another immovable object to prevent it from slipping away from the participants while testing.

PROCEDURE

Have participants remove their shoes and sit down at the test apparatus with the knees fully extended and the feet shoulder-width apart. The feet should be flat against the board. The arms are extended forward with the hands placed on top of each other to perform the test. Without bouncing or jerking, participants reach directly forward, palms down, along the measuring scale four times. The position of maximum reach on the fourth trial should be held for one second.



Test position for the sit and reach test

SCORING AND ANALYSIS

The score is the most distant point reached on the fourth trial measured to the nearest centimeter. This point must be touched by the fingertips of *both* hands. Refer to tables for percentile ranking. Scores below the 50th percentile represent poor extensibility in the lower trunk areas: posterior thigh, low back, or posterior hip.

The location of musculature tightness may be determined by observing participants as they perform the test. A poor performance that is coupled with a "rounding" of both the upper and lower back suggests tightness in the posterior thigh. Poor performance accompanied by a "rounding" of the upper back and a relatively "straight" lower back suggests a tightness in the lower back musculature.

It should be noted that many preadolescent boys and girls might *not* be able to reach the 23 centimeter level. This is because the legs may become proportionately longer in relation to the trunk during a growth spurt.

RELIABILITY AND VALIDITY

The sit and reach test has been validated against several other types of flexibility tests. The coefficients obtained have generally ranged between .80 and .90. Reliability coefficients for the sit and reach test have been high, ranging above .70 (AAHPERD, 1980).

Sit-Up and Reach (cm) – BOYS†

Percentile	Age												
	5	6	7	8	9	10	11	12	13	14	15	16	17+
99	36	37	38	38	37	37	38	52	41	43	47	45	48
75	29	29	28	29	29	28	29	29	30	33	34	36	40
50	25	26	25	25	25	25	25	26	26	28	30	30	34
25	22	22	22	22	22	20	21	21	20	23	24	25	28
5	17	16	16	16	16	12	12	13	12	15	13	11	15

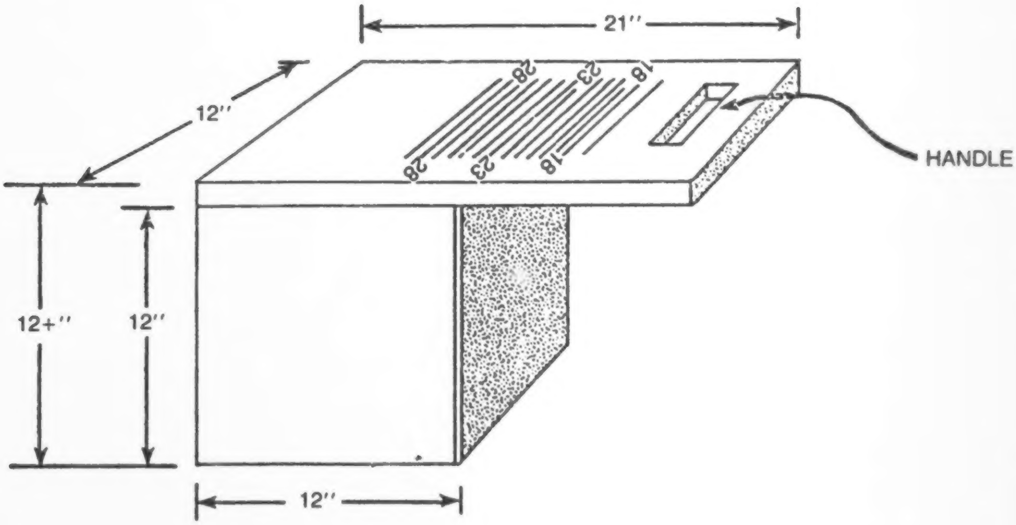
Sit-Up and Reach (cm) – GIRLS†

Percentile	Age												
	5	6	7	8	9	10	11	12	13	14	15	16	17+
99	37	38	37	39	39	41	41	46	49	49	49	48	47
75	30	30	31	31	31	31	32	34	36	38	41	39	40
50	27	27	27	28	28	28	29	30	31	33	36	34	35
25	23	23	24	23	23	24	24	25	24	28	31	30	31
5	18	18	16	17	17	16	16	15	17	18	19	14	22

---

†Reprinted by permission of the American Alliance for Health, Physical Education, Recreation and Dance, 1900 Association Drive, Reston, Virginia 22091.

SIT AND REACH BOX



Schematic drawing of the sit and reach measurement apparatus

1. Using any sturdy wood or comparable construction material (3/4 inch plywood seems to work well), cut the following pieces:

   2 pieces - 12 in x 12 in
   2 pieces - 12 in x 10 1/2 in
   1 piece - 12 in x 21 in

2. Assemble the pieces using nails or screws and wood glue.
3. Inscribe the top panel with one centimeter gradations. It is crucial that the 23 centimeter line be exactly in line with the vertical panel against which the subject's feet will be placed.
4. Cover the apparatus with two coats of polyurethane sealer or shellac.
5. For convenience, a handle can be made by cutting a 1 inch x 3 inch hole in the top panel.
6. The measuring scale should extend from about 9 to about 50 cm.

References

American Alliance for Health, Physical Education, Recreation, and Dance. (1980). *Health related physical fitness test manual*. Reston, VA: Author.

American Alliance for Health, Physical Education, Recreation, and Dane. (1984). *Technical manual: Health related physical fitness*. Reston, VA: Author.

American Heart Association. (1972). *Exercise testing and training of apparently healthy individuals: A handbook for physicians*. Dallas, TX: Author.

Astrand, P.O., & Ryhming, I. (1954). A homogram for calculation of aerobic capacity (physical fitness) from pulse rate during submaximal work. *Journal of Applied Physiology*, *7*(1), 218-221.

Blair, S.N., Falls, H.B., & Pate, R.R. (1983). A new physical fitness test. *The Physician and Sportsmedicine*, *11*(4), 87-95.

Bruce, R.A., Kusumi, F., & Hosmer, D. (1973). Fundamentals of clinical cardiology: Maximal oxygen intake and nomographic assessment of functional aerobic impairment in cardiovascular disease. *American Heart Journal*, *85*(4), 546-562.

Cooper, K.H. (1977). *The aerobics way*. New York, NY: Bantam.

Golding, L.A., Myers, C.R., & Sinning, W.E. (Eds.). (1982). *The Y's way to physical fitness*. Chicago, IL: YMCA of the USA.

Kasch, F.W., & Boyer, J.L. (1968). *Adult fitness, principles and practice*. Palo Alto: Mayfield Publishing.

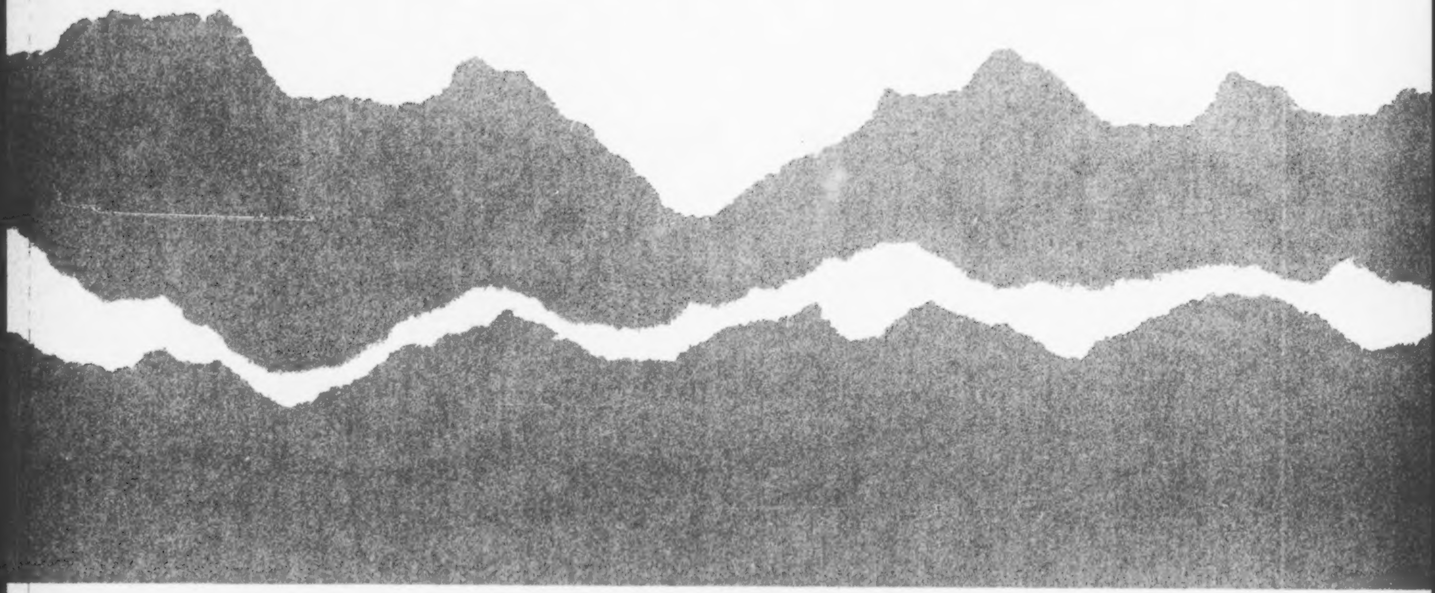
Kasch, F.W., Phillips, W.H., Ross, W.D., Carter, J.E.L., & Boyer, J.L. (1966). A comparison of maximal oxygen uptake by treadmill and step-test procedures. *Journal of Applied Physio*logy, *21*(4), 1387-1388.

Lohman, T.G., Roche, A.F., & Martorell, R. (Eds). (1988). *Anthropometric standardization reference manual*. Champaign, IL: Human Kinetics Books.

Malina, R.M., Hamill, P.V.V., & Lemesow,S. (1973). *Selected body measurements of children 6-11 years: United States* (Vital and Health Statistics, Series 11, No. 123). Washington, D.C.: U.S. Government Printing Office.

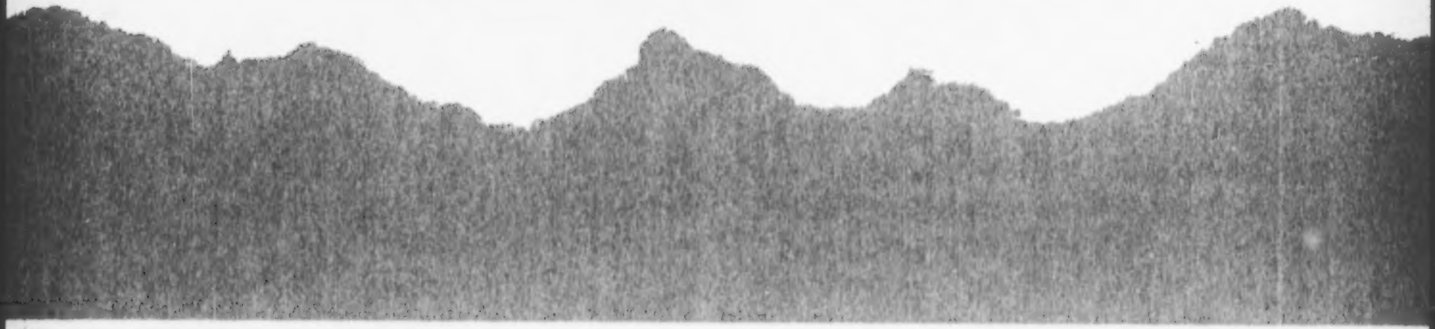
Pate, R.R. (1983). A new definition of fitness. *The Physician and Sportsmedicine*, *11*(4), 77-82.

Pollock, M.L., Wilmore, H.J., & Fox, S.M. (1984). Exercise in health and disease: Evaluation and prescription for prevention and rehabilitation. Philadelphia, PA: W.B. Saunders.



CHAPTER FIVE

Locally Conducted Psychometric Studies



Locally Conducted Psychometric Studies

As described in Chapter One, the first step in using the newly developed handbook measures to examine program effectiveness is to select those that match program goals. However, evaluators cannot assume that a measure that appears to assess a desired program outcome will produce valid data about that outcome. When evaluators use a measure, they first want to determine the technical quality of that measure to ensure that any conclusions drawn about a program's effects are warranted. The purpose of this chapter is to assist evaluators in conducting validation studies for those handbook measures chosen for use in program evaluation.

Determining the Technical Quality of Measuring Devices

The degree to which a measuring instrument yields scores from which one can make legitimate inferences is referred to as validity. Tests are not valid or invalid. Rather, it is the inferences made, based on test results, that are valid or invalid. It is, therefore, technically accurate to focus on the *validity of score-based inferences* rather than the validity of a particular measuring device.

The concept of validity is highly dependent on the particular way in which a measuring instrument will be used. For example, a measure of the use of injury prevention skills may permit a valid inference regarding the *number* of different skills that program participants use, but may yield invalid inferences regarding the *frequency* with which participants use each skill. Furthermore, a test may yield valid inferences for a particular purpose with one population but invalid inferences for the same purpose with a different population. Thus, because validity varies on the basis of purpose and population, it is most appropriate to examine validity in the setting in which a measure will be used.

A second factor in determining the technical quality of a measurement instrument deals with the extent to which the instrument produces reliable, that is, consistent, results. Because the newly developed handbook measures have been subjected only to small-scale field tests, no reliability data are currently available. It is hoped that handbook users will conduct their own reliability studies and share those results with the Centers for Disease Control. In this way, results can be compiled over time and, subsequently, provided to handbook users. Procedures for evaluating the reliability of the handbook measures will be presented following a discussion of local validation approaches.

Categories of Validity Evidence

There are three major types of evidence regarding validity. These include content-related evidence of validity, criterion-related evidence of validity, and construct-related evidence of validity. The procedures for securing each type of validity evidence will be described below.

Content-related evidence of validity. Content-related evidence of validity involves the careful review of a measure's content by individuals identified as experts in the content area being assessed. This type of validity evidence is particularly important for measures designed to assess examinees' knowledge and skills. To secure positive content-related validity, the measure must include only those items that correspond to the content area being assessed, and its items must address all important facets of that content area. The systematic, expertise-rooted procedures used to develop the handbook's instruments helped to ensure that appropriate content was built into the measures. Subsequent reviews by external experts confirmed that the measures are, indeed, focused on suitable content. These development procedures and the role of expert advisors in the project are described in the handbook's preface.

If there are questions regarding the suitability of the content in any of the handbook's measures, content-related validity can be examined by assembling a panel of experts who can judge the suitability of a measure's content for the specific program evaluation purpose for which the measure is to be used. A panel of approximately 10 knowledgeable individuals can be asked to review the measuring instrument's items, one by one, and render independent yes/no judgments regarding the appropriateness of each item's content (in relationship to the inference that the program evaluators wish to make on the basis of the measure). In addition, panelists can be asked to determine whether any important content has been omitted from the measure. For example, if a knowledge measure such as **Facts About Exercise** is being reviewed, panelists might be asked first to think of all the important facts about exercise that program participants must know and then to indicate the percentage of those facts that are present in the measure being reviewed. This straightforward indication of a measure's content representativeness, when coupled with judgments regarding the content appropriateness of a measure's items, can yield important content-related evidence of validity for a measure.*

Criterion-related evidence of validity. Criterion-related evidence of validity requires that a measure be checked against an independent criterion. The independent criterion or standard should be one that the measure would be expected to predict. Criterion-related validity is most important for the handbook measures in the areas of behavior and intention. In the area of behavioral self-reports, for example, criterion-related validity would focus on the degree to which the self-reports reflect actual behavior. So, for example, criterion-related validity for a self-report instrument designed to measure individuals' exercise patterns would be secured by correlating responses on this instrument with observations (by others) of the extent to which exercise was *actually* being done.

External criterion measures, such as observations, while often more accurate measures of behavior than self-reports, are extremely costly and time consuming to use. Thus, although it may be possible to use such criterion measures in a one-time validity study, they typically will not eliminate the need for self-report instruments in routine program evaluations. The general procedure for conducting a criterion-related validity study is shown in Figure 5.1.

A correlation of approximately .50 or higher between the measure and criterion would indicate that the new measure is predictive of the external criterion measure and, therefore,

---

* For additional information about how to conduct content-related validation studies, see Annotated Bibliography Nos. 18, 23, 27, and 34.

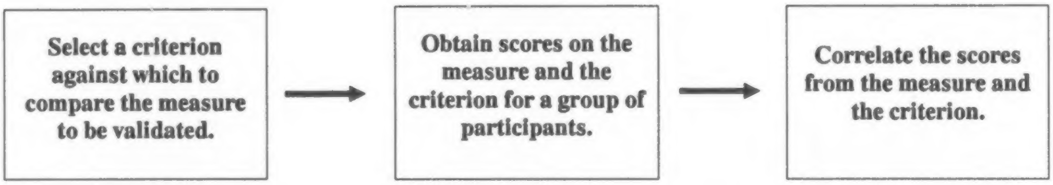

Select a criterion against which to compare the measure to be validated.
Obtain scores on the measure and the criterion for a group of participants.
Correlate the scores from the measure and the criterion.

Figure 5.1: Procedure for conducting criterion-related validity studies

is measuring what it is intended to measure. A low correlation would call into question the self-report instrument as a measure of the behavior of interest.

Each criterion-related validity study must be specifically designed for the particular measure being examined and the purpose for which it will be used. For example, imagine that an evaluator wanted to examine the criterion-related evidence of validity for the handbook's measure entitled **Intention to Exercise**. The evaluator must first identify an appropriate criterion measure. How is a program evaluator likely to use an intention measure? The most likely use would be to employ it as a proxy measure foreshadowing a program's effect on the future behavior of participants. That is, will program participants begin or continue to exercise in the future? Thus, an appropriate criterion measure might be reported physical activity several months following the program.

To assemble criterion-related evidence of validity for the intention measure, a program evaluator could administer the intention measure at the end of the program to a group of at least 30 participants (or repeat this process each session until responses from at least 30 participants are obtained) and obtain completed self-report surveys several months later regarding participants' physical activity. Once both measures are collected for every individual, a correlation could be computed between the strength of intention for exercising regularly and whether regular exercise was being done following the program. Thus, the criterion-related validity study would examine whether the intention measure was, in fact, predictive of later behavior. A measure that can serve as a meaningful proxy for participants' future behavior can prove highly useful in the evaluation of a program's impact on participants.*

Construct-related evidence of validity. The final type of validity evidence to be reviewed, construct-related evidence of validity, is particularly important for those handbook measures that do not have a clear criterion against which they can be evaluated. Such measures include the attitudinal and affective measures such as **Exercising Regularly**, a measure that examines an individual's perceived ability to exercise in different situations. Construct-related validity involves the gradual accumulation of data regarding what a test measures. Three strategies are customarily used to secure construct-related evidence of validity for a measure. First, in the *related-measures strategy,* predictions can be tested about

---

* For additional information about the design and analysis of criterion-related validity studies, see Annotated Bibliography Nos. 18, 23, 27, and 34.

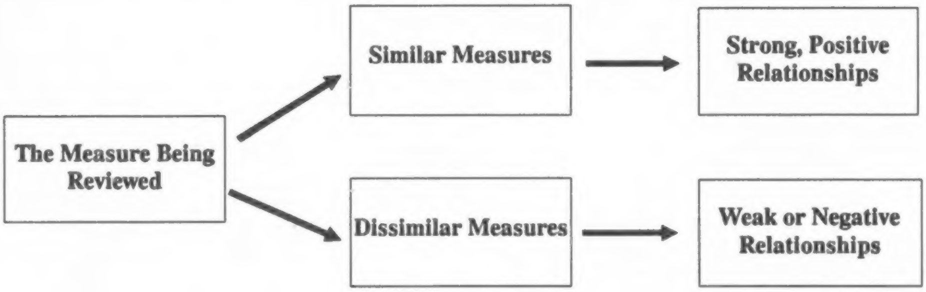

Similar Measures
Strong, Positive Relationships
The Measure Being Reviewed
Dissimilar Measures
Weak or Negative Relationships

Figure 5.2: Correlations between measures assessing similar/dissimilar attitudinal dimensions

the extent to which the measure of interest is correlated with other measures. For example, perceived ability to exercise should be positively related to other measures aimed at assessing a similar attribute but should show reduced correlations with measures tapping different attitudinal dimensions. Thus, other existing measures can be correlated with the measure of interest to help clarify what is being measured.

If the correlations are consistent with the prior predictions, then construct-related evidence of validity has been obtained to support the defensibility of inferences based on the measure's use. Figure 5.2 illustrates the anticipated correlations between the measure of interest and other similar and dissimilar measures.

A second approach to examining construct-related validity involves predictions about group differences and is referred to as a *differential-populations strategy*. For this procedure, two or more groups are identified that are expected, based on other characteristics, to perform differently on the measure of interest. For example, the two groups might consist of individuals who exercise regularly versus those who do not. If the anticipated performance difference between the two groups is not obtained, it would raise the question as to whether the test was measuring what it was thought to measure.

A third strategy for securing construct-related evidence of validity is referred to as an *intervention strategy* because it involves the use of interventions such as training programs. For instance, a measure examined via this strategy could be administered to a group of participants before and after a "proven" physical fitness training program. If a difference in participants' scores on the measure is not observed, then the construct-related evidence of validity regarding the measure being reviewed is not supportive of the measure's use.

Construct-related evidence of validity is never based on a single study. Instead, consideration of a variety of studies, employing multiple validation strategies such as those described here, will help provide greater clarification regarding the appropriateness of using a given measuring instrument.*

Types of Reliability

A second characteristic of a defensible measurement instrument is the reliability or consistency with which it measures. The reliability of a test can be examined in three distinct ways. These include test-retest reliability, alternate-forms reliability, and internal consistency. Each of these approaches will be described below.

Test-retest reliability. Test-retest reliability (also referred to as *stability* reliability) examines the extent to which a measurement instrument is consistent over testing occasions. That is, will an individual who received a particular score on one testing occasion receive a similar score on a different testing occasion? Typically, to secure test-retest reliability information, an instrument is administered once to a group of individuals (30 or more). The same instrument is then administered again under similar conditions to the same group of individuals approximately two to four weeks later. Individuals' scores from the two administrations are then correlated. The higher the correlation, the greater the stability of measurement over time. Short tests, or other tests that are likely to be easily remembered, may result in an overestimate of reliability if participants recall their answers and, hence, respond similarly on the second testing occasion.

Alternate-forms reliability. The knowledge and skill measures in this handbook have two forms that may be used for a pretest to posttest comparison. The administration of one form for the pretest and the other form for the posttest is desirable because the pretest may sensitize participants to pay more attention to those issues included on the pretest than to other equally important issues. However, to draw defensible conclusions based on the use of two different forms at pretest and posttest, the forms must be equivalent.

To examine alternate-forms reliability, it is necessary to administer both forms to the same group of individuals. The scores from the two forms can then be correlated. High correlations indicate that the same conclusions would be drawn about an individual or group of participants regardless of which of the two forms had been used. Thus, there would be reliable or consistent measurement across alternate forms. A high alternate-forms reliability coefficient does not guarantee that the forms are perfectly equidifficult. If the two forms are not of equal difficulty, that is, participants perform consistently better on one form than the other, it would still be possible to obtain high between-forms correlations. Thus, it is important to be attentive to mean scores on the two test forms. It is also permissible to use p-values (the percentage of examinees getting each item correct) to reassign items to forms so that they are more equidifficult. After the redistribution of items, a second alternate-forms reliability study should be conducted.

---

* For additional information about how to conduct construct-related validity studies, see Annotated Bibliography Nos. 18, 23, 27, and 34.

Handbook users should not assume equivalence or equidifficulty for the multiple forms provided in this handbook. Until alternate-forms reliability and test difficulty are examined, the measures should be used in a design such that half of the participants take Form A as a pretest and Form B as a posttest while the other half take Form B as a pretest and Form A as a posttest. This counterbalancing technique eliminates the possible influence of one form being more difficult than the other.

Internal consistency. Internal consistency examines the extent to which the instrument measures a single or related set of constructs. The higher the internal consistency, the greater the homogeneity of items on the test. A test thought to measure a single attitudinal dimension should have relatively high internal consistency reliability. Procedures for calculating internal consistency include split-half reliability, Kuder-Richardson formulas, and Cronbach's Alpha. The split-half reliability coefficient is calculated by administering the test to a group of at least 30 participants and then correlating scores from the odd versus the even items. A correction for test length must then be made using the Spearman-Brown formula. The split-half procedure is very similar to alternate-forms reliability in that two "forms" are correlated by separating the odd and even items. Kuder-Richardson formulas for internal consistency provide an estimate of the average of all possible split-halves. These formulas, like Spearman-Brown, require that test items be binary-scored, that is, able to be scored as right or wrong. Cronbach's Alpha is identical to Kuder-Richardson for binary-scored items but can also be used for items that yield responses to which several points can be assigned, such as the items on **Effects of Exercise**.

Not all forms of reliability need to be computed for every test. For example, alternate-forms reliability would be computed only for those measures that have two forms. Internal consistency estimates are less appropriate for multidimensional measures. Test-retest reliability is appropriate for most measures but often presents pragmatic problems due to the need to retest the same individuals.*

Groups and Individuals

The validity and reliability procedures reviewed here were originally developed to examine the quality of tests used for *individual* assessment purposes. In contrast, the recommended use of the handbook measures is to perform *group* analyses for program evaluation. Thus, the appropriate reliability issue is whether scores for a group of individuals are relatively consistent. Similarly, the validity issue is whether changes in scores for a group of individuals are reflective of changes in the group's knowledge, skills, affect, or behavior. Because group scores are more stable than individual scores, the procedures outlined above are likely to underestimate the reliability and validity of the measures when used for program evaluation. Practically speaking, a measurement instrument with a lower reliability or validity coefficient would be acceptable when used for group rather than individual diagnosis. For example, Salvia and Ysseldyke (1981, p. 98) have recommended the following minimum standards for alternate-forms reliability:

---

* For additional information about how to examine the reliability of measurement instruments, see Annotated Bibliography Nos. 3, 18, 19, 23, 27, and 34.

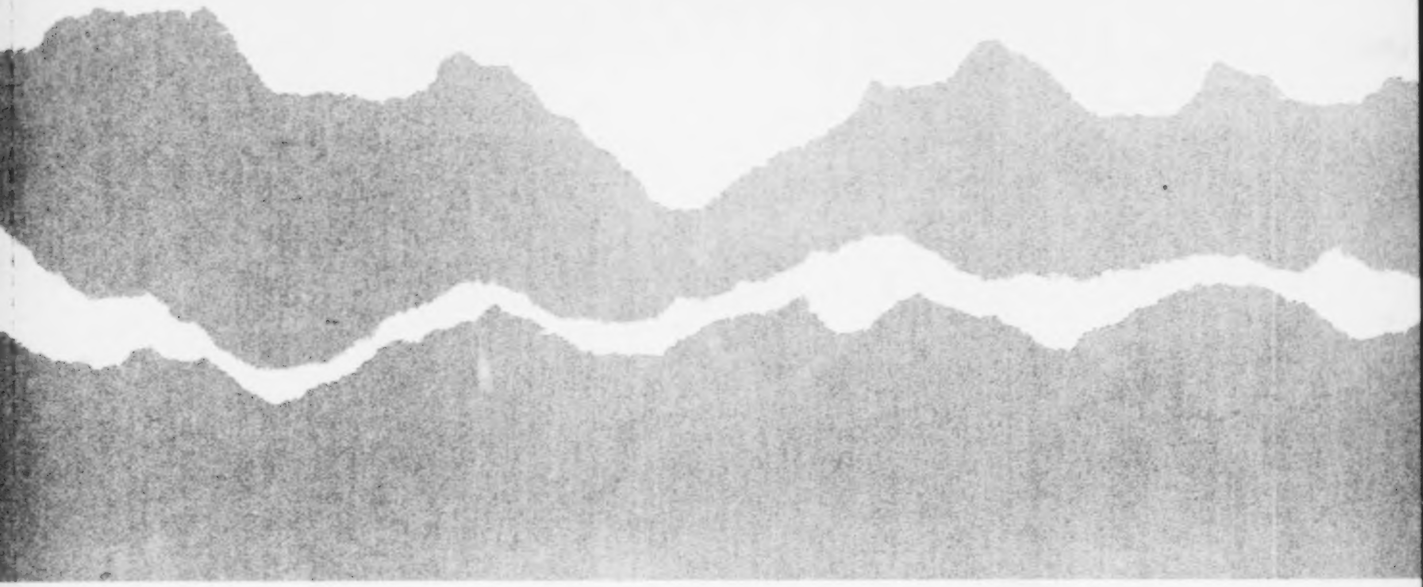
.60 - when scores are reported for groups

.80 - when scores are used for individual screening

.90 - when scores are used for important educational decisions for individuals

Thus, standards for acceptable reliability and validity vary depending on the purpose for using a particular measure. However, minimal levels for each are critical for making sound decisions about a program. With a little creativity and effort, studies of reliability and validity can often be integrated into the ongoing operation of a program.

In addition to providing a brief overview, the major purpose of this chapter was to encourage handbook users to conduct local reliability and validity studies and to consider the involvement of a measurement specialist or the use of appropriate references in designing such studies. As suggested at the outset of the chapter, if such local studies are carried out, results should be forwarded to the Centers for Disease Control (Attention: Dr. Diane Orenstein, Project Officer, Center for Health Promotion and Education, Centers for Disease Control, 1600 Clifton Road N.E., Atlanta, GA 30333). This information will be shared with future handbook users.



Appendices

Appendix A

AMPLIFIED CONTENT DESCRIPTORS

PLANNING TO EXERCISE SAFELY

(Adult/Adolescent Measure)

PLANNING A SAFE EXERCISE PROGRAM

(Preadolescent Measure)

Program Design/Activity Selection

1. Choosing an exercise program without proper guidance or participating in an activity program where everyone exercises at the same level of effort can be dangerous or ineffective.
2. A person choosing activities for an exercise program should consider the following factors: health status, physical fitness level, age, desired benefits, preference for individual or team activities, preference for indoor activities, as well as available time, money, and equipment.
3. A person does not need to be athletic to participate in an exercise program.
4. Heart/lung endurance activities are the most important part of any health-focused exercise program.
5. Different sports and activities provide different benefits.
6. People who are beginning to exercise should not exercise as long and hard as people who exercise regularly.
7. The benefits of exercise depend upon the type of activity as well as the intensity, duration, and frequency of exercise.
8. People who exercise regularly will reach a point where their fitness level will not improve unless they exercise harder, longer, and/or more often.
9. Medicine does not improve a person's exercise ability.

Avoiding Injury

10. On smoggy days, people should exercise less than they usually do.
11. It is better to run on soft surfaces such as dirt than hard surfaces such as concrete sidewalks.
12. Much of the heat built up by the body during play or exercise is released as sweat.
13. Heat, dampness in the air, and lack of wind or clouds are factors that can lead to heat problems.

14. A person should drink plenty of liquids before, during, and after playing or exercising in the heat.

15. Heat exhaustion is the first sign that the body is becoming less able to control its temperature.

16. Some signs of heat exhaustion are weakness, breathlessness, dizziness, rapid heart beat, and a body temperature of around 101°F.

17. Cold stress can occur when the body begins to lose heat faster than it can be built up.

18. Cold stress is a result of exercising in very low temperatures and is made worse by wind or damp air.

19. Exercising or playing in very cold weather can cause frostbite.

20. Pain in the joints, feet, ankles, or legs mean that a person should stop or reduce exercising or playing.

21. Following the rules of a game, using equipment the right way, and learning the right ways to move during exercise and play can help a person avoid getting hurt.

22. A person should call a doctor right away if there is pain in the left side or middle of the chest, in the left side of the neck, or in the left shoulder or arm during or just after exercising.

23. People who have health problems or are over 30 years old and not used to exercising hard should to see a doctor before starting an exercise program.

Exercise Program Components

24. An exercise program should include a warm-up and stretching period, an endurance period, a cool-down period, and, if desired, a strength conditioning period.

25. A complete physical fitness program should include heart/lung, flexibility, muscular strength, and muscular endurance exercises.

26. People of all ages should warm up and stretch before exercising and cool down after exercising.

27. The warm-up and stretching period of an exercise session may consist of easy stretching and calisthenics and should last from 5 to 15 minutes.

28. Warm-up and stretching exercises improve flexibility and get the body ready to do heart/lung exercises.

29. The warm-up and stretching period before exercising and the cool-down period after exercising will prevent or help decrease muscle and joint soreness.

30. Cool-down exercises usually consist of easy jogging or walking and light stretching and should last from 5 to 10 minutes.

Heart/Lung Fitness

31. Heart/lung endurance is the ability of the body to continue rhythmic exercise for a long period of time and to efficiently make and distribute energy.

32. In order to improve heart/lung fitness, heart/lung exercises must be done at least three times per week.

33. To build heart/lung endurance, people should exercise at their target heart rate, rather than as hard as they can.

34. A person's target heart rate is the number of heart beats per minute that should be reached during exercise in order to get the most heart/lung benefit.

35. The target level for a healthy person building heart/lung endurance is 60 to 85 percent of one's maximum heart rate.

36. Heart/lung endurance exercises should keep the heart rate at its target level for at least 20 minutes.

37. The difficulty of heart/lung fitness exercises can be checked by measuring the heart rate.

38. Everyone does not have the same target heart rate.

39. Exercising at 85 percent or more of one's maximum heart rate may be dangerous unless a person is in excellent physical condition.

40. Brisk walking, skipping rope, running, uphill hiking, swimming, bicycling, cross-country skiing, ice hockey, and rowing will improve one's heart/lung endurance.

41. If played hard and without stopping, handball, racquetball, squash, badminton, tennis, and basketball will improve one's heart/lung endurance.

42. Baseball, bowling, football, golf, softball, and volleyball will not improve one's heart/lung endurance.

Muscular Endurance and Strength Conditioning

43. Muscular endurance is the ability of one muscle or a muscle group to exercise for a long time.

44. Muscular endurance can be improved by repeatedly doing exercises that require strength.

45. Muscular strength is the ability of one muscle group to use force or move a heavy weight one time.

46. Muscular strength can be improved through activities that require muscle groups to exercise against heavy resistance.

47. Bicycling, handball, racquetball, squash, running, skating, skiing, swimming, tennis, walking, and calisthenics, can improve muscular strength and endurance to some extent.

48. People should do muscular strength and endurance activities at least two days a week to maintain their muscular strength and endurance.

49. Flexibility is the ability to move one joint in the body through its full range of motion, that is, all the possible ways it should be able to move.

Flexibility Conditioning

50. The safest way to improve flexibility is by doing activities that require a joint to move slowly through its full range of motion.
51. Calisthenics, judo, karate, swimming, and yoga can improve muscular flexibility.
52. People should do flexibility exercises every day to keep their bodies flexible.

FACTS ABOUT EXERCISE

(Adult/Adolescent Measure)

EXERCISE FACTS

(Preadolescent Measure)

Exercise and Fitness

1. In order to promote cardiorespiratory fitness and to improve health, exercise must involve the heart muscle and other parts of the cardiorespiratory system.
2. A person's physical fitness level refers to how well a person's heart, lungs, muscles, and other organs work, not to how the body looks.
3. A person must exercise regularly in order to be physically fit.
4. Without exercise, the body begins to weaken and has a greater chance of developing certain diseases.
5. Experts agree that regular, moderate exercise is a good form of preventive medicine.
6. In general, middle-aged and older people benefit from regular exercise as much as young people do.
7. A regular exercise program does not need to take a great deal of time.

Effects on the Cardiorespiratory System

8. Regular exercise can increase the number of blood vessels that go to the skeletal muscles and increase the amount of blood they can carry to certain parts of the body.
9. Regular exercise can improve the circulation of blood to the heart and the skeletal muscles.
10. Regular exercise can make the heart pump more blood with each heartbeat.
11. Regular exercise can reduce the level of cholesterol in the blood, thereby keeping the blood vessel linings clear of fatty materials.
12. Regular exercise can improve stroke volume, which is the amount of blood pumped with each beat of the heart.
13. Regular exercise can decrease the resting blood pressure.
14. Regular exercise reduces the heart rate at rest and allows an individual to exercise at near maximum rates for longer periods.
15. Regular exercise improves the efficiency of breathing so that fewer breaths are needed to get the same volume of air.
16. Regular exercise increases the maximum amount of oxygen that the body can process while exercising.
17. Regular exercise can slow down the natural decline in lung capacity that normally occurs with age or inactivity.

18. The cardiovascular improvement effects of exercise are generally greatest at the beginning of a fitness program.
19. Regular exercise can help keep the lungs healthy as a person gets older.
20. If the frequency, intensity, and/or duration of an exercise program are progressively increased, cardiovascular improvement may continue for several years.
21. All the cardiovascular benefits that result from regular exercise are gradually lost if exercise is not continued.
22. A single exercise session will have no lasting effect on the cardiovascular system.
23. The cardiovascular benefits of exercise occur only when the exercise program is of sufficient intensity to raise the heart rate to at least 60% of its maximum rate.
24. Physically inactive individuals are at greater risk of developing coronary heart disease and high blood pressure than are active individuals.
25. Physically inactive individuals are more likely than active individuals to have a heart attack.
26. Physically inactive individuals are more likely than active individuals to die immediately after a heart attack.
27. Regular exercise may be prescribed to help individuals with asthma and emphysema.
28. Pulse or heart rate at rest is an accepted measure of cardiovascular fitness for most people.
29. Pulse or heart rate can be measured at either the wrist or at one of the blood vessels in the neck.
30. The effects of exercise on the respiratory system are determined by measuring the amount of air a person breathes to perform a standard amount of exercise.
31. The maximum amount of oxygen the body can burn during hard exercise is called maximum oxygen consumption (VO_2 max).
32. Maximum oxygen consumption levels are used as a measure of cardiorespiratory fitness.

Effect on the Muscular/Skeletal System

33. Regular exercise can firm muscles, restore their tone, and increase muscular strength and flexibility.
34. Muscles that are regularly exercised are better able to use fat to produce energy.
35. Regular exercise can reduce minor muscular aches, pains, stiffness, and soreness.
36. Regular exercise can build the muscular strength and endurance necessary to carry on normal daily activities easily and efficiently.
37. Regular exercise can strengthen the bones, ligaments, and tendons.
38. Muscles that are not exercised become smaller, but do not turn into fat.
39. People who are muscularly fit are not necessarily cardiovascularly fit.

40. Regular exercise may help prevent and relieve lower back muscle pain.
41. Regular exercise is sometimes prescribed by doctors to help problems due to arthritis.
42. Regular exercise can correct some problems in posture.
43. Good posture as a child can help keep a person from having back pains as an adult.

Body Composition and Weight Reduction

44. Body composition is defined as the amount of fat compared to the amount of lean body weight.
45. The percent of fat is a better indicator of appropriate body composition than is total body weight.
46. Lean body weight refers to the amount of non-fat tissue a person has.
47. Exercise increases the number of calories used by an individual.
48. The body will continue to use calories at an increased rate for a short time after exercise has stopped.
49. Regular exercise can increase the body's ability to mobilize and use fat.
50. Regular exercise can help reduce excess body fat and total body weight.
51. Weight loss through regular exercise maximizes fat loss and minimizes protein loss.
52. There is little or no evidence that localized exercise or spot reducing can reduce fat in isolated areas of the body.
53. Regular exercise combined with dieting is more effective than dieting alone for reducing body weight.
54. The number of calories burned during exercise depends on the type of activity, the intensity and duration of the activity, and an individual's body weight.
55. A heavier person uses more calories and fat than a lighter person during a comparable exercise period.
56. The rate of weight loss may be estimated by determining caloric intake and caloric expenditure based on the frequency, duration, and intensity of exercise.
57. Underwater weighing and skinfold measurement can be used to determine an individual's body composition.

Psychological and Sociological Effects

58. Exercise is a socially acceptable way of reducing tension.
59. Experts believe that regular exercise can reduce anxiety, stress, and depression.
60. Experts believe that regular exercise is likely to increase a person's self-assertiveness and feelings of self-reliance.

61. Studies show that there is a positive relationship between physical fitness, mental alertness, and emotional well-being.

62. Regular exercise can improve an individual's self-image.

63. Physical fitness has been positively associated with improved work performance and reduced time away from work or school.

64. Exercise can provide the opportunity to meet new friends and spend time with family and friends.

65. Participation in a regular exercise program can improve one's ability to fall asleep quickly and to sleep well.

Physical Fitness Promotion Content Bibliography

American Red Cross. (1979). *Standard first aid and personal safety* (2nd ed.). Garden City, NY: Doubleday.

Antonacci, R.J., & Barr, J. (1975). *Physical fitness for young champions*. New York: McGraw-Hill.

Cooper, K. (1968). *Aerobics*. New York: Evans and Co.

Craig, M. (1973). *Miss Craig's growing-up exercises*. New York: Random House.

Folkins, C.H., & Sime, W.E. (1981). Physical fitness training and mental health. *American Psychologist*, 36(4), 373-389.

Golding, L.A., Myers, C.R., & Sinning, W.E. (Eds.). (1982). *The Y's way to physical fitness*. Chicago, IL: National Board of YMCA.

Grawunder, R., & Steinmann, M. (1980). *Life and health* (3rd ed.). New York: Random House.

Morehouse, L.E., & Miller, A.T., Jr. (1967). *Physiology of exercise* (5th ed.). St. Louis: The C.V. Mosby Co.

Pollock, M.L., Wilmore, J.H., & Fox, S.M. (1984). *Exercise in health and disease: Evaluation and prescription for prevention and rehabilitation*. Philadelphia: W.B. Saunders Co.

President's Council on Physical Fitness. (1980). *Adult physical fitness*. Washington, DC: U.S. Government Printing Office.

Sharkey, B.J. (1979). *Physiology of fitness: Prescribing exercise for fitness, weight control, and health.* Champaign, IL: Human Kinetics Publishers.

Strauss, R. (Ed.). (1979). *Sports medicine and physiology*. Philadelphia: W.B. Saunders Co.

U.S. Department of Health and Human Services. (1981). *Exercise and your heart*. (National Institute of Health Publication No. 81-1677). Washington, DC: U.S. Government Printing Office.

Wilmore, J.H. (1977). *Athletic training and physical fitness: Physiological principles and practices of the conditioning process*. Boston: Allyn & Bacon.

Wilmore, J.H. (1981). *The Wilmore fitness program: A personalized guide to total fitness and health*. New York: Simon and Schuster.

Appendix B

INFORMED CONSENT PROCEDURES

Prior to administering measures to participants, program personnel should inform participants about the content covered by the measures and the purpose of the program's evaluation study. Program personnel may also wish to provide the opportunity for participants to indicate whether or not they consent to participate in the study and complete the selected measures. Informed consent is obtained by presenting all information pertinent to the study and asking the participant to affix a signature indicating that the information has been read and that consent is given to participate.

If the decision is made to obtain informed consent, program personnel have the choice of employing a "passive" consent procedure or an "active" consent procedure. *Passive informed consent* consists of asking participants to sign and return a consent form only *if they do not wish to participate* in the study. Participants who do not return the consent form are considered eligible to participate in the study.

Active informed consent requires participants to sign and return the consent form *if they wish to participate.* Only those participants who return a signed form can be included in the study. Consequently, the participation rate resulting from an active consent procedure is generally lower than that obtained from a passive consent procedure.

To construct an informed consent form, program personnel should consider including the following items:

1. A general statement of the program goals and objectives.
2. A brief explanation of the study procedures and measures.
3. An indication that the participant is free to withdraw consent and to discontinue participation at any time.
4. An explanation of the procedures to be taken to ensure anonymity and confidentiality of responses.
5. An indication that participants are free not to answer specific items or questions.
6. A place for the participants to affix their signatures under a statement indicating that the participant agrees to participate (active consent) or does not agree to participate (passive consent) in the study. If appropriate, a date for the return of the consent form should be specified.

Appendix C

ANNOTATED EVALUATION BIBLIOGRAPHY

1. Alkin, M.C., & Solmon, L.C. (Eds.). (1983). *The costs of evaluation*. Beverly Hills, CA: Sage.

   In this collection of essays both theoretical and practical issues relevant to cost-focused program evaluations are presented.

2. American Psychological Association. (1973). *Ethical principles in the conduct of research with human participants*. Washington, DC: Author.

   This treatise focuses on the appropriateness of carrying out various types of research investigations with human subjects. Because the American Psychological Association has had a long-standing concern about ethical issues in the conduct of research investigations, this publication will be of interest to numerous evaluators of health education programs.

3. American Psychological Association, American Educational Research Association, National Council on Measurement in Education. (1985). *Standards for educational and psychological tests*. Washington, DC: Author.

   This volume presents the most widely used set of standards for psychological and educational tests. Frequently cited by users of educational tests, the standards have recently been employed in numerous judicial deliberations. Relatively brief, the standards should be consulted by health educators who employ assessment devices regularly.

4. Anderson, L.W. (1981). *Assessing affective characteristics in the schools*. Boston: Allyn and Bacon.

   Anderson provides an excellent set of practical suggestions for the creation of affective assessment instruments. He includes one of the most easily understood expositions of various scaling procedures including Likert, Thurstone, and Guttman scales.

5. Bausell, R.B. (Ed.). *Evaluation and the health professions*. Newbury Park, CA: Sage.

   This quarterly publication deals with a variety of evaluation-relevant issues of interest to health educators.

6. Berk, R.A. (Ed.). (1982). *Handbook of methods for detecting test bias*. Baltimore: The Johns Hopkins University Press.

   This collection of individual essays offers the reader a comprehensive depiction of methods currently available to detect the presence of bias in tests.

7. Berk, R.A. (Ed.). (1984). *A guide to criterion-referenced test construction*. Baltimore: The Johns Hopkins University Press.

   This collection of essays consists of papers presented at the first Johns Hopkins University National Symposium on Educational Research. In addition, a number of more recently written chapters have been included in this revision of a 1980 text. The authors address many of the important problems, both conceptual and technical, facing developers and users of criterion-referenced measures.

8. Campbell, D.T., & Stanley, J.C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

   This volume, originally a chapter in a larger volume, has had substantial impact on the fields of research and evaluation. Evaluators of health education programs will wish to consider this truly classic treatment of data-gathering designs suitable for experimental and quasi-experimental settings.

9. Churchill, G.A., Jr. (1979). *Marketing research: Methodological foundations* (2nd ed.). Hinsdale, IL: The Dryden Press.

   Although written in the context of marketing research, this textbook covers several topics of vital importance in evaluation. Topics such as research design, data collection, sampling, and data analysis are covered in a readily understandable yet accurate way. An excellent resource.

10. Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (rev. ed.). New York: Academic Press.

    Cohen offers a useful treatment of factors which should be considered when one draws samples for use in research or evaluation activities. Of special interest is the set of easy-to-use guidelines he offers for determining the estimated sample size necessary to detect differences between groups.

11. Cook, T.D., & Campbell, D.T. (1976). The design and conduct of quasi-experiments and true experiments in field settings. In M.D. Dunnette (Ed.), *Handbook of industrial and organizational psychology*. Chicago: Rand McNally.

    This is an updated version of the famous exposition of quasi-experimental and experimental data-gathering designs by Donald T. Campbell and Julian C. Stanley (see Reference No. 8). An excellent discussion of four types of validity is featured in this essay.

12. Cook, T.D., & Campbell, D.T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.

    This widely cited volume provides a comprehensive treatment of quasi-experimental investigations in settings of substantial relevance to the concerns of health educators. There are excellent discussions of internal and external validity, including the various threats to both types of validity. A systematic consideration of the commonly used data-gathering designs is offered, including an extended appraisal of interrupted time-series designs.

13. Cordray, D.S., Bloom, H.S., & Light, R.J. (Eds.). (1987, Summer). *Evaluation practice in review* (New Directions for Program Evaluation, No. 34). San Francisco: Jossey-Bass.

    This volume contains a set of thought-provoking chapters dealing with what has been learned about the practice of evaluation during the past decade. The chapters on evaluation politics by Eleanor Chelimsky and on naturalistic evaluation by Egon Guba would be of particular interest to evaluators of health education programs.

14. Cronbach, L.J. (1963). Course improvement through evaluation. *Teachers College Record, 64*, 672-683.

    This article is an early piece, presenting the virtues of what would later be termed "formative" evaluation. It rings as true today as it did more than two decades ago, and it applies as much to evaluation in health education as it does to more traditional evaluation. Emphasizing the role of evaluation in gathering information that can improve programs, this article is well worth reading.

15. Cronbach, L.J. (1977). *Analysis of covariance in nonrandomized experiments: Parameters affecting bias.* Unpublished occasional paper, Stanford Evaluation Consortium, Stanford University.

    A highly technical piece on the complications associated with using analysis of covariance, this article is recommended only for those prepared to handle a critical data-analysis problem in a sophisticated way.

16. Cronbach, L.J., Ambron, S.R., Dornbusch, S.M., Hess, R.D., Hornik, R.C., Phillips, D.C., Walker, D.F., & Weiner, S.S. (1980). *Toward reform of program evaluation.* San Francisco: Jossey-Bass.

    This important book considers the function of evaluation in a pluralistic society and presents 95 theses on the role of evaluators and evaluations. In addition to providing a contemporary conception of evaluation, it provides a historical and multidisciplinary perspective of the field. This volume will be of considerable interest to those evaluating health education programs.

17. Cronbach, L.J., & Furby, L. (1970). How should we measure 'change'—or should we? *Psychological Bulletin, 74*, 68-80.

    A technical treatise on the dangers associated with using gain scores. A very significant piece, but recommended only for those with some psychometric training.

18. Cunningham, G.K. (1986). *Educational and psychological measurement.* New York: Macmillan.

    This is a standard introductory text focusing on the major topics associated with measurement as it applies to such tasks as program evaluation.

19. Ebel, R.L. (1979). *Essentials of educational measurement* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

    This is a standard, easily read introductory text, covering important topics in the field of educational testing. Ebel, a prominent leader of traditional educational testing practices, provides a lucid treatment of a wide range of measurement topics.

20. Fetterman, D.M., & Pitman, M.A. (Eds.). (1986). *Educational evaluation: Ethnography in theory, practice, and politics*. Beverly Hills, CA: Sage.

    This collection of essays touches on ethnographically oriented evaluation of educational programs. Health educators wishing to learn about this recently emphasized approach to educational evaluation will find this volume of interest.

21. Green, L.W. (1979). Research methods translatable to the practice setting: From rigor to reality and back. In S.J. Cohen (Ed.), *New directions in patient compliance* (pp.141-151). Lexington, MA: Lexington Books.

    Green attends to a practical dilemma facing those who evaluate health education programs, namely, the necessity to make trade-offs between validity and feasibility in field settings. Six strategies for coping with evaluation under adverse circumstances are described.

22. Green, L.W., & Figa-Talamanca, I. (1974). Suggested designs for evaluation of patient education programs. *Health Education Monographs, 2* (1), 54-71.

    In this essay Green and Figa-Talamanca suggest data-gathering designs for conducting evaluations of patient education programs. The authors also explore several issues related to evaluations of this variety.

23. Green, L.W., & Lewis, F.M. (1986). *Measurement and evaluation in health education and health promotion*. Palo Alto, CA: Mayfield.

    This volume is an excellent resource for health educators concerned with the evaluation of their programs. Green and Lewis provide a series of useful explanations of topics in both measurement and health evaluation. Their expositions are peppered with practical examples drawn from health education and health promotion.

24. Hambleton, R.K., Swaminathan, H., Algina, J., & Coulson, D.B. (1978). Criterion-referenced testing and measurement: A review of technical issues and development. *Review of Educational Research, 48* (1), 1-48.

    This is a comprehensive review of the field of criterion-referenced testing. Hambleton and his colleagues do a masterful job of isolating the key issues in criterion-referenced testing and describing results of research investigations bearing on those issues. Somewhat technical at times, this review is one of the more widely cited essays dealing with criterion-referenced testing.

25. Hays, W.L. (1973). *Statistics for the social sciences*. New York: Holt, Rinehart, and Winston.

    This comprehensive text handles basic and advanced statistical considerations. Somewhat technical at points, Hays nonetheless provides an excellent set of step-by-step guidelines to statistical practice.

26. Joint Committee on Standards for Educational Evaluation. (1981). *Standards for evaluations of educational programs, projects, and materials*. New York: McGraw-Hill.

The development of these evaluation standards was spearheaded by a joint committee of the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education. Thirty standards are presented, addressing issues related to deciding whether to evaluate, defining the evaluation problem, designing the evaluation, budgeting for the evaluation, collecting and analyzing data, and reporting the evaluation. Intended for both consumers of evaluation and individuals conducting evaluations, this reference may be of most use to evaluators who are relatively new to the field.

27. Kubiszyn, T., & Borich, G. (1987). *Educational testing and measurement: Classroom application and practice* (2nd ed.). Glenview, IL: Scott-Foresman.

Another introductory text dealing with the nuts and bolts of measurement, this book will provide health educators with a good overview of educational measurement.

28. Levin, H.M. (1975). Cost-effectiveness analysis in evaluation research. In M. Guttentag & E.L. Struening (Eds.), *Handbook of evaluation research* (Vol. 2, pp. 89-122). Beverly Hills, CA: Sage.

This essay probes the important considerations involved in determining cost-effectiveness of programs in the context of educational evaluations. Theoretical as well as practical guidelines are provided.

29. Levin, H.M. (1983). *Cost-effectiveness: A primer* (New Perspectives in Evaluation, Vol. 4). Beverly Hills, CA: Sage.

This text is a splendid introduction to the fundamental concepts of cost analysis on program evaluation. Levin provides succinct descriptions along with advantages and disadvantages for cost-feasibility, cost-effectiveness, cost-benefit, and cost-utility analyses.

30. Linn, R.L., & Slinde, J.A. (1977). The determination of the significance of change between pre- and posttesting periods. *Review of Educational Research, 47*, 121-150.

This article reviews many of the major issues in the measurement of change from pretesting to posttesting periods and suggests possible alternatives. These authors share the general sentiment of many others in the field that "more is expected from gain scores than they can reasonably be expected to provide."

31. Lord, F.H. (1963). Elementary models for measuring change. In C.W. Harris (Ed.), *Problems in measuring change* (pp. 21-38). Madison: Wisconsin Press.

This is an early treatise on the problems associated with measuring change. Although this chapter rapidly becomes very technical, the early sections provide an intuitive explanation of the difficulties with using gain scores.

32. Mark, M.M., & Shotland, R.L. (Eds.). (1987, Fall). *Multiple methods in program evaluation* (New Directions for Program Evaluation, No. 35). San Francisco: Jossey-Bass.

    Decrying the infrequency with which multiple methods are used in program evaluation, six chapters are offered in this volume, not only advocating multiple methods, but also describing how such program evaluations can be conducted.

33. Oakland, T. (Ed.). (1977). *Psychological and educational assessment of minority children.* New York: Brunner/Mazel.

    This collection of essays provides a series of useful suggestions for those who are more sensitive to the possible bias present in educational tests.

34. Popham, W.J. (1981). *Modern educational measurement.* Englewood Cliffs, NJ: Prentice-Hall.

    Varied topics in the field of educational measurement are introduced in this text. Norm-referenced measurement and criterion-referenced measurement are both considered, with the special applications of criterion-referenced assessment emphasized. Chapters on the relationship of testing to teaching and the measurement of affect will be of special interest to health educators.

35. Popham, W.J. (1988). *Educational evaluation.* Englewood Cliffs, NJ: Prentice-Hall.

    This is an introductory text, written in fairly nontechnical language, about the field of educational evaluation. Evaluators of health education programs will find it simple to translate the book's contents to their own specialties.

36. Popham, W.J., & Sirotnik, K.A. (1973). *Educational statistics: Use and interpretation* (2nd ed.). New York: Harper and Row.

    This easily read introductory text deals with the fundamental types of statistical considerations needed by program evaluators. It is intended for those who are not particularly comfortable with mathematical approaches to statistics.

37. Riecken, H.W., & Boruch, R.F. (1971). *Social experimentation: A method for planning and evaluating social intervention*. New York: Academic Press.

    This is a significant contribution to our thinking about large-scale social interventions, their design and appraisal. It provides a useful analysis of the ways that the experimental method can be defensibly employed in connection with major social programs.

38. Rivlin, A.M., & Timpane, P.M. (Eds.). (1975). *Ethical and legal issues in social experimentation*. Washington, DC: Brookings Institution.

    Rivlin and Timpane explore the sorts of legal and ethical issues to which evaluators of health education programs must attend.

39. *SPSS-X User's Guide* (3rd ed.). (1988). Chicago: SPSS Inc.

    This is a widely used, well-organized set of "canned" computer analysis programs for use in the social sciences. Health educators who have occasion to use computer analyses will find the SPSS manual most helpful.

40. Salvia, J., & Ysseldyke, J.E. (1981). *Assessment in special and remedial education* (2nd ed.). Boston: Houghton Mifflin.

This text, intended for individuals who must apply assessment to special education and remedial education, provides measurement insights for health educators who deal with such populations of learners.

41. Scriven, M. (1967). The methodology of evaluation. In R.W. Tyler, R.M. Gagné, & M. Scriven (Eds.). *Perspectives of curriculum evaluation* (pp. 39-83). Chicago: Rand McNally.

This seminal article was the first essay in which Scriven distinguished between the now commonly accepted formative and summative roles of evaluators. Scriven addresses a wide variety of topics, emphasizing the importance of comparative appraisals of two or more programs' merits.

42. Scriven, M. (1972). Prose and cons about goal-free evaluation. *Evaluation Comment, 3*, 1-4.

In this essay Scriven offers goal-free evaluation as an antidote to excessive preoccupation with the program staff's expressed objectives. Scriven argues that evaluators should attend to the results produced by a program, not the rhetoric of its program goals.

43. Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill.

This is the classic treatment of nonparametric statistical techniques. Although a bit out of date these days, Siegel's text offers the most easily understood treatment of nonparametric statistical procedures. Because of the author's admitted zealousness in support of nonparametric techniques, those using Siegel's text should also consult a critique of it by Robert Savage, *Journal of American Statistical Association*, 1957, *52*, 331-344.

44. Suchman, E.A. (1967). *Evaluative research: Principles and practice in public service and social action programs.* New York: Russell Sage Foundation.

In this volume, Suchman provides extensive coverage of the application of the experimental research model in conducting evaluations. Although evaluation has come a long way since this book was written, the volume provides a clear description of the predominant conceptualization of evaluation in the past decade.

45. Tukey, J.W. (1977). *Exploratory data analyses.* Reading, MA: Addison-Wesley.

Creative approaches to displaying and understanding data are provided by Tukey in this excellent demystification of data analysis.

46. Walberg, H.J., Postlethwaite, T.N., Creemers, B.P.M., & de Court, E. (Eds.). (1987). Educational evaluation: The state of the field. *International Journal of Educational Research, 11* (1).

This special issue, as its title suggests, presents comprehensive review of field of program evaluation from authors based in the U.S. and abroad.

47. Webb, E.J., Campbell, D.T., Schwartz, R.D., Sechrest, L., & Grove, J.B. (1981). *Nonreactive measures in the social sciences* (2nd ed.). Dallas: Houghton Mifflin.

This charming volume provides readers with a series of powerful and clever tactics to secure data, particularly of an affective nature, without sensitizing respondents to the evaluator's purposes.

48. Weiss, C.H. (1972). *Evaluation research: Methods of assessing program effectiveness.* Englewood Cliffs, NJ: Prentice-Hall.

Weiss offers a pithy overview of prominent program evaluation considerations including the formulation of questions to be addressed, the design of the evaluation study, and the utilization of evaluation results. A paperback, this brief book (160 pp.) offers an excellent introduction to what Weiss refers to as "evaluation research."

49. Windsor, R.A., Baranowski, T., Clark, N., & Cutter, G. (1984). *Evaluation of health promotion and education programs.* Palo Alto, CA: Mayfield.

This text is a useful introduction to the evaluation of health education programs. Windsor et al. have provided readers with a series of health-relevant examples to illustrate their explorations.

50. Worthen, B.R., & Sanders, J.R. (Eds.). (1973). *Educational evaluation: Theory and practice*. Worthington, OH: C.A. Jones.

This volume was one of the earliest compilations of various program evaluation models applied to education. Evaluation theorists whose views are presented in this book include Stake, Cronbach, Scriven, Tyler and others. Worthen and Sanders have authored sections of the book and have included a series of original chapters by a number of evaluation specialists. While focused on educational evaluation in general, the volume is of substantial relevance to program evaluation of health education programs.

51. Worthen, B.R., & Sanders, J.R. (1987). *Educational evaluation: Alternative approaches and practical guidelines.* New York: Longman.

This introductory text is organized around a series of alternative approaches to educational evaluation, including the "objectives-oriented" and "advisory-oriented" approaches.

52. Worthen, B.R., & White, K.R. (1987). *Evaluating educational and social programs: Guidelines for proposal review, onsite evaluation, evaluation contracts, and technical assistance.* Boston: Kluwer-Nijhoff.

This volume provides a first-rate series of practical guidelines dealing with varied aspects of proposal review, onsite evaluation, evaluation contracts, and technical assistance.

53. Zdep, S.M., & Rhodes, I.N. (1977). Making the randomized response technique work. *The Public Opinion Quarterly, 40*, 531-537.

This easily read essay describes the randomized response technique, a procedure used to obtain sensitive information from respondents more accurately than if respondents were directly asked about sensitive information.